

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 5, 2006, 23:14:07 ; Search time 437 Seconds
(without alignments)
6551.022 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 MSSNPFLEKGFVRLDPE.....RHEMFLQTNVYKPGMSFA 1020

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abes/ABSSWEB_spool/US10645746/runat_05072006_110149_5129/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs808 -USER=US10645746 @CGN 1.1 271 @runat_05072006_110149_5129 -NCPU=6
-ICPU=3 -NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	577	10.8	7478	3	US-09-949-016-948 Sequence 948, App
2	577	10.8	7478	3	US-09-949-016-4453 Sequence 4453, App
3	575.5	10.8	2910	3	US-09-533-029-105 Sequence 105, App
4	452.5	8.5	1662	3	US-09-270-767-11928 Sequence 11928, A
5	388	7.3	1874	3	US-09-270-767-12369 Sequence 12369, A
6	312	5.8	2328	3	US-10-043-774B-1 Sequence 1, Appli
7	306	5.7	3472	3	US-09-873-737A-5 Sequence 5, Appli
8	304	5.7	4064	3	US-09-873-737A-3 Sequence 3, Appli

9	292	5.5	3620	4	US-10-094-749-902	Sequence 902, App
10	288	5.4	2579	3	US-10-104-047-1469	Sequence 1469, Ap
11	278.5	5.2	3138	3	US-10-104-047-1676	Sequence 1676, Ap
12	265	5.0	771	3	US-09-248-796A-6819	Sequence 6819, Ap
13	256.5	4.8	734	4	US-09-297-648-2528	Sequence 2528, Ap
14	205	3.8	3047	3	US-09-873-737A-1	Sequence 1, Appli
15	178.5	3.3	2167	3	US-09-270-767-13282	Sequence 13282, A
16	170	3.2	300	4	US-09-297-648-1917	Sequence 1917, Ap
17	167	3.1	9626	3	US-09-150-867-2	Sequence 2, Appli
18	167	3.1	9626	5	US-09-724-584-2	Sequence 2, Appli
19	165.5	3.1	1664976	3	US-08-916-421B-1	Sequence 1, Appli
20	165.5	3.1	1664976	3	US-09-692-570-1	Sequence 1, Appli
21	164.5	3.1	8503	3	US-09-620-312D-130	Sequence 130, App
22	163	3.0	6276	3	US-09-949-016-2009	Sequence 2009, Ap
23	163	3.0	6284	3	US-09-949-016-1028	Sequence 1028, Ap
24	162.5	3.0	5192	3	US-09-949-016-5831	Sequence 5831, Ap
25	162.5	3.0	6773	3	US-09-166-350-27	Sequence 27, Appl
26	161.5	3.0	8257	3	US-09-595-684B-30	Sequence 30, Appl
27	158	3.0	8530	3	US-09-949-016-5562	Sequence 5562, Ap
28	158	3.0	10300	3	US-09-949-016-636	Sequence 636, App
29	158	3.0	38575	3	US-09-949-016-17304	Sequence 17304, A
30	158	3.0	119153	3	US-09-949-016-12378	Sequence 12378, A
31	157	2.9	10136	2	US-08-353-700-2	Sequence 2, Appli
32	157	2.9	10136	7	PCT-US95-16216-2	Sequence 2, Appli
33	156	2.9	4868	2	US-08-139-937-12	Sequence 12, Appl
34	156	2.9	4868	7	PCT-US93-11310-12	Sequence 12, Appl
35	156	2.9	8789	2	US-08-328-254-5	Sequence 5, Appli
36	155	2.9	6921	3	US-09-643-597-117	Sequence 117, App
37	155	2.9	6921	3	US-09-480-884A-117	Sequence 117, App
38	155	2.9	6921	3	US-09-542-615A-117	Sequence 117, App
39	155	2.9	6921	3	US-09-506-421B-117	Sequence 117, App
40	155	2.9	6921	3	US-09-221-107-117	Sequence 117, App
41	155	2.9	6921	3	US-09-466-396A-117	Sequence 117, App
42	155	2.9	6921	3	US-09-476-496A-117	Sequence 117, App
43	155	2.9	6921	3	US-09-630-940B-117	Sequence 117, App
44	155	2.9	6921	3	US-09-285-479-117	Sequence 117, App
45	155	2.9	6921	3	US-10-007-700-117	Sequence 117, App

ALIGNMENTS

RESULT 1
US-09-949-016-948
Sequence 948, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 948
LENGTH: 7478
TYPE: DNA
ORGANISM: Human
US-09-949-016-948

Alignment Scores: 1.17e-55 Length: 7478
Pred. No.: 577.00 Matches: 237
Score: 577.00 Conservative: 173
Percent Similarity: 40.2% Mismatches: 370
Best Local Similarity: 23.2% Indels: 243
Query Match: 10.8% Gaps: 38
DB: 3

US-10-645-746-3 (1-1020) x US-09-949-016-948 (1-7478)

QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu----- 37
DB 209 ATGGGATGAA-GCGGGACCTCGGAGCAGCTGCGGCGCTTACCTGCCCCCTGCAG 267
QY 38 -----LysLysValLeuLeu 42
DB 268 CAGGTGTTCCAGCACCTCGCGGCTGGCATTTGGCACTGTGGGGAACCAATCAAGCTC 327
QY 43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGlu 62
DB 328 CTGGCCAAATACTTTGAGGTGGACATCCCTAAGATCGAGCTGTACCACTACGAGGTGGAC 387
QY 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82
DB 388 ATCAAGCCGGATAAG-----TGTCCTCCGCTAGAGTC 417
QY 83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
DB 418 AACCGGGAAGTGGTGA-----TACATGGTCCAGCATTTCAAGCTC 459
QY 103 GlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLysAspThrValTyr 120
DB 460 CAG-----ATCTTGGTGATCGCAAGCCTGTGTATGATCGAAGGAAGAACATTATAC 510
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal----- 138
DB 511 -----ACTGTACAGCACTGCCCATTTGGCAACGCAAGCGGTGCACTTT 552
QY 139 ---ValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTyr 157
DB 553 GAGGTGACAAATCCCTGGGAAGGAAGGAT---CGAATCTTTAAGGTCTCCATCAAGTGG 609
QY 158 ThrMetIleLeuTyrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
DB 610 CTAGCATTTGTAGCTGGCGAATGCTGCATGAGGCCCTGGTCAGCGGCCAGATCCCTGTT 669
QY 178 LysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysVal 197
DB 670 CCTTGGAGTCTGTGCAAGCCCTGGATGTGGCATGAGGCATGAGGCATCC-----ATG 723
QY 198 ArgTyrAlaProPheValAsnGluIleLysValGlnPheAlaLysAsnPheValTyr 217
DB 724 AGGTACACCCCTGTG----- 738
QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
DB 739 ---GGCGGCTCTTCTTCTCACCGCCTGAGGGCTACTACCACCG----- 780
QY 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGlu 257
DB 781 -----CTGGGGGGTGGCGCGAAGTCTGGTTCGGCTTTCACCAAGTCTGTGGCCCT 831
QY 258 LeuPheAspGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
DB 832 GCCATG---TGAAGATGATGCTCAACATGTATGCTCAGCCACTGCCCTTTTATAGGCA 888
QY 278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAsp 297
DB 889 -----CAGCCAGTGATTCATGTGTGAGGTCTGCACATCAGGAACAATAGATGAG 942
QY 298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
DB 943 CAGCCCAAG-----CCCTCAGGACTCT 966
QY 318 AlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLys-----CysAla 334
DB 967 CAGCGGTTCGCTTCCCAAGAGATCAAGGGCTCAAGGTGAGAGTCAACCACTGTGGA 1026
QY 335 GluValTrpAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr 350
DB 350 -----GluMetSerArgLeuThrGluArgHisLeuThr 350

DB 1027 CAGATGAAGAGGAGTACCGCGTGTGTATGTTATCCCGTCCGCTCTAGCCATCAGACA 1086
QY 351 PheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArg 370
DB 1087 TTCCCTCTTACAGCTGGAGAGTGGACAGACTGTG----- 1119
QY 371 GlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGluAsnLysLys 390
DB 1120 -----GAGTGCACAGTGGCCACATATTTCAAGCAGAAATATAAC 1158
QY 391 Phe-----IleGluPheProHisLeuProLeuValLysValLysSerGlyLysGluTyr 409
DB 1159 CTTACAGTCAAGTATCCCATCTGCGCTTACAAGTTGGCCAGCAACAAAGCATACC 1218
QY 410 AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIle 429
DB 1219 TACCTTCCCTAGAGGTCTGTAAACATT---GTGGCTGGCAGCGCTGTATTAAAGAGCTG 1275
QY 430 AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys 449
DB 1276 ACCGACACACAGACCTCGACCATGATTAAGGCCACAGTAGATCCGCTCCAGACACAG 1335
QY 450 GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal 469
DB 1336 GAGGAGATCAGTCGCGCTGATGAAGATGCCAGCTACAACCTAGATCCC-----TACATC 1389
QY 470 GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
DB 1390 CAGGAATTTGGATCAAAATGAAGATGACATGACGAGGTGACAGGCGGAGTGTGCGC 1449
QY 490 GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro----- 505
DB 1450 GCGCCCATCTTCAGTACCGCGCGCGGAACCGGGCCATTCGCCACCCCAATCAGGCTGTC 1509
QY 506 ---ValIleArgGlyPheGlnGluLysGlnLeu---AsnValValProGluLysGluLeu 523
DB 1510 TGGGACATGTCGGGG-----AAACAGTTCTACATGGGATGAGATCAAAAGCTGG 1560
QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
DB 1561 GCCATCGCTGCTTCGCCACCCCAAAA-----CAGTGTGCGAAGAG----- 1602
QY 544 ValValLysPheTyrThrGluLeuIleGly-----GlyCysLysPhe 557
DB 1603 GTGCTCAAGAACTTCACAGCAGCTGCGGAAGATTTCCAAGATGCGGGGATGCTCATC 1662
QY 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
DB 1663 CAGGGTCAACCTTGTCTTCTGCAAAATATGCACAGGGGCA-GACAGCGTG----- 1710
QY 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPhe 597
DB 1710 ----- 1710
QY 597 eGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGln 617
DB 1711 -----GAGCTATGTTCCGGCATCTCAAGAACACACCTACTCAGG 1748
QY 617 sValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHI 637
DB 1749 GCTGAGCTCAATTAATTTGTCATCTCCAGGAAGACGCGGTGTATGCTGAGGTGAACG 1808
QY 637 eTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
DB 1809 TGTCCGAGATACACTCTTGGGAATGGCTACCGAGTGTGTGAGGTGAAGACGTTGTC 1868
QY 657 sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAla 677
DB 1869 GACC-----TCACCTCAGACTCTGTCCAACTCTGCTCT 1901
QY 677 uLysIleAsnAlaLysLysLeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAla 697
DB 1902 CAAGATCAATGTCAAACTTGGTGGCATTAACAACATCTCTA----- 1941

```

QY 697 uileSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyI 717
Db 1942 -----GTCCACACAGCGCTCTGCGCTTTTCAACAGCCAGTG---ATATTCCTGGGAGC 1994
QY 717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAl 737
Db 1995 AGATGTTACACACCCACAGCGGATGGGAAACCTTCTATCACAGCATGGGTAGG 2054
QY 737 aSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysAr 757
Db 2055 CAGTATGGATGCCACCCAGCGCATCTGCTACTGTGCGGTACAG-----2103
QY 757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
Db 2104 -----CGACCACCGCAAGAGATCATTTGAAGACTTGTGTC 2135
QY 775 aLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794
Db 2136 CTACATGGTGGTGGAGCTCTCTCATCCAAATCTACAGTCCACCGGTTTCAG---CCTAC 2192
QY 794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
Db 2193 CCGCATCATCTTCTACCGAGATGGGTGCTGAAGGCCAGCTACCCAGATACCTCCATT 2252
QY 814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834
Db 2253 TGAGCTACTGGCCATTCGTGATCGCTCATCAAACTG-----GAAAGGA 2297
QY 834 pProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuAr 854
Db 2298 CTACACGCTGGGATCATTTATATTGTGGTGGAAGAACGCCATCACACCGCTTTTC-- 2355
QY 854 gArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspVa 874
Db 2356 -TGTCGTGACAAAGATGAGCGAATTGGGAAGAGTGGTAACATCCAGCT-----2403
QY 874 lAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyI 894
Db 2403 -----2403
QY 894 eValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAs 914
Db 2404 -----GGGACCAAGTGGGACCAACATCACCCACCATTTGAGTTGA 2447
QY 914 pPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrVa 934
Db 2448 TTCTATCTGTGAGCCAGCGATCCAGGCGCAGCCGACCATCCCATTTACTATGT 2507
QY 934 lMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAl 954
Db 2508 TCTTTGGGATGCAACCGTTTTCACAGCAGATGAGTCCAGATCCTGACGTACCGCTGTG 2567
QY 954 aPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHi 974
Db 2568 CCACACTTACGTACGATGACACGCTGTCTCTATCCAGACCTGCCCTACTATGCCCG 2627
QY 974 sLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAs 994
Db 2628 CCTGGTGGCTTTCCGGGCAGCA---TACCACCTGGTGGACAGGAGCATCAGCTGGAGA 2684
QY 994 P 994
Db 2685 G 2685

```

RESULT 2

```

US-09-949-016-4453
; Sequence 4453, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4453
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4453

```

Alignment Scores:

```

Pred. No.: 1,17e-55 Length: 7478
Score: 577.00 Matches: 237
Percent Similarity: 40.2% Conservative: 173
Best Local Similarity: 23.2% Mismatches: 370
Query Match: 10.8% Indels: 243
DB: Gaps: 38

```

US-10-645-746-3 (1-1020) x US-09-949-016-4453 (1-7478)

```

QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu-----37
Db 209 ATGGGATGGA-CGGGACCTCGGAGCAGCTGGGGCGCTTACCTGCCCCCTCGAG 267
QY 38 -----LysLysValLeuLeu 42
Db 268 CAGGTGTTCCAGCACCTCGCGCGCTGGCATTGGCATTGGGAAACCAATCAAGCTC 327
QY 43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGlu 62
Db 328 CTGGCCAAATTAATTTTCTGAGGTGGGACATCCCTTAAGATCGACGCTGTACCACTACGAGGTGGAC 387
QY 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82
Db 388 ATCAACCGCGATGAG-----TGTCCCGTAGAGTC 417
QY 83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
Db 418 AACCGGGAAGTGGTGAA-----TACATGCTCCAGCATTTCAAGCCT 459
QY 103 GlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLysAspThrValTyr 120
Db 460 CAG-----ATCTTTGGTATCGCAAGCCTGTGTATGATCGAAGAAAGAACATTTAC 510
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal-----138
Db 511 -----ACTGTACAGCACTCCCATTTGGCAACGACGGGTGCGATTT 552
QY 139 ---ValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTyr 157
Db 553 GAGGTGACAATCCTCGGGAAGGAGGAT---CGAATCTTTAAGGTCTCCATCAAGTGG 609
QY 158 ThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
Db 610 CTAGCCATTGTGAGCTGGCGAATGTGCATGAGGCCCTGTGTGCGGCCAGATCCCTGTT 669
QY 178 LysAspGluGluAlaAsnArgSerTyrLysPheLysAsnValMetThrGlnLysVal 197
Db 670 CCTTGGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGCACCTGGCATCC-----ATG 723
QY 198 ArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTyr 217
Db 724 AGGTACACCCCTGTG-----738
QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237

```

Db 739 ---GGCGCTCTCTTCTCACCGCTGAGGGCTACTACACCG----- 780
Qy 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGlu 257
Db 781 ---CTGGGGGTGGCGCGAGGCTGGTTCGGCTTTCACCAAGCTGTGTGGCCCT 831
Qy 258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
Db 832 GCATG---TGGAGATGATGCTCAACATGTCTCAGCCACTCCCTTTTATAGGCA 888
Qy 278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuValAspProGlnSerCysAsnAsp 297
Db 889 ---CAGCCAGTGATTCAGTTTCATGTGTGAGGTCTGCACATCAGCAACATAGATGAG 942
Qy 298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
Db 943 CAGCCCAAG-----CCCTCAGCGACTCT 966
Qy 318 AlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys-----CysAla 334
Db 967 CAGCGGTTTCGTTTACCAGGAGATCAAGGGCTGAGGTGGAGTCAACCACTGTGA 1026
Qy 335 GluValTrpAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr 350
Db 1027 CAGATGAGAGGAAGTACCGCGTGTGTAATGTTACCGCTGCGCTGTAGCCATCAGACA 1086
Qy 351 PheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArg 370
Db 1087 TTCCCTTACAGCTGGAGAGTGGACAGACTGTG----- 1119
Qy 371 GlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGluAsnLysLys 390
Db 1120 -----GAGTGCACAGTGGCCACATATTTCAAGCAGAAATATAAC 1158
Qy 391 Phe----IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyr 409
Db 1159 CTTGAGTCAAGTATCCCATCTGCCCTGTACAAAGTTGGCCAGGAACAAAGCATACC 1218
Qy 410 AlaValProMetGluHisLeuGluValHisGluLysProGluArgTyrLysAsnArgIle 429
Db 1219 TACCTTCCCTAGAGCTGTGAACATT---GTGGCTGGCGAGCGCTGTATTAAAGCTG 1275
Qy 430 AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys 449
Db 1276 ACCGACAACAGACCTCGACCATGATAAGGCCACAGCTAGATCCGCTCCACAGACAG 1335
Qy 450 GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal 469
Db 1336 GAGGAGATCAGTCGCGCTGATGAAGATGCCAGCTACAACTTAGATCCC-----TACATC 1389
Qy 470 GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
Db 1390 CAGGAATTTGGGATCAAAAGTAGGATGATGATGACGGAGGTGACAGGGCGAGTGTGCGG 1449
Qy 490 GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro----- 505
Db 1450 GCGCCCATCTTCAGCTACGCGCGCGGAACCGGCCATTGCCACACCCCAATCAGGGTGTG 1509
Qy 506 ----ValIleArgGlyPheGlnGluLysGlnLeu---AsnValValProGluLysGluLeu 523
Db 1510 TGGGACATCGGGGG-----AAACAGTTCTACAAATGGGATGAGATCAAAAGTCTGG 1560
Qy 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
Db 1561 GCCATGCGCTGTCTGCACCCCAAAAA-----CAGTGTGAGAAAG----- 1602
Qy 544 ValValLysPheTyrThrGluLeuIle-Gly-----GlyCysLysPhe 557
Db 1603 GTGCTCAAGAACTTACACAGCAGCTGCGGAAGATTTCCAGAGATGCGGGGATGCTATC 1662
Qy 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
Db 1663 CAGGGTCAACCTTGTCTTCTGCAAAATATGACAGGGGGCA-GACAGCGTG----- 1710

Qy 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPh 597
Db 1710 ----- 1710
Qy 597 eGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLys 617
Db 1711 -----GAGCCTATGTTCCGGCATCTTCAAGAACACCTACTCAGG 1748
Qy 617 sValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis 637
Db 1749 GCTCAGAGCTCATATTGTTCATCTCTCCAGGAAGACCGCGGTGTATGCTCAGGTGAACG 1808
Qy 637 sTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
Db 1809 TGTCGGAGATACACTCTTGGGAATGGCTACGCGAGTGTGTGAGGTGAAGAACGTGTCTCA 1868
Qy 657 sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLe 677
Db 1869 GACC-----TCACCTCAGACTCTGTCCAACCTCTGCCT 1901
Qy 677 uLysIleAsnAlaLysLeuGlyIleAsnGlnLeuLeuAspTrpSerGluIleAlaGln 697
Db 1902 CAAGATCAATGTCAAACTTGGTGGCATTAACAACATCTTA----- 1941
Qy 697 uIleSerProGluGluLysGluArgLysThrMetProLeuThrMetTyrValGlyIle 717
Db 1942 ----GTCCACACACCGCGCTCTGCGTTTTTCAACAGCCAGTG---ATATTCTGGAGC 1994
Qy 717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAl 737
Db 1995 AGATGTTACACACCCCGCAGCAGCGGATGGGAAAAAACCTTCTATCACAGCAGTGTAGG 2054
Qy 737 sSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysAr 757
Db 2055 CAGTATGATGTCACCCCGCAGCGCATCTGTCTACTGTGCGGGTACAG----- 2103
Qy 757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
Db 2104 -----CGACCAGCGCAGAGATCATTTGAAGACTGTGC 2135
Qy 775 aLysPheValLys----LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794
Db 2136 CTACATGTGCGTGAGCTCTCATCAATTCTACAAGTCCACCCGTTTCAAG---CCTAC 2192
Qy 794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
Db 2193 CGGCATCATCTTCTACCGAGATGGGGTGCTGTAAGGCCAGCTACCCAGATATCTCCACTA 2252
Qy 814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834
Db 2253 TGAGTACTGGCCATCTGTGATGCTGCATCAAACTG-----GAAAAGGA 2297
Qy 834 pProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuAr 854
Db 2298 CTACCAGCTGGGATCACTTATTTGTGTGTCAGAAACGCCATCACCCCGCTTTTC-- 2355
Qy 854 gArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspVa 874
Db 2356 -TGTCCTGACAGAATGAGCGAATTTGGGAAGAGTGGTAACATCCCGAGCT- 2403
Qy 874 lAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIle 894
Db 2403 ----- 2403
Qy 894 eValAsnProSerSerGlyThrValAspLysLeuIleValSerLysTyrLysPheAs 914
Db 2404 -----GGACCCAGTGGACACCAACATCACCCACCCATTTGAGTTGA 2447
Qy 914 pPhePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrVa 934
Db 2448 CTTTCTATCTGTGACCCAGCGAGGCATCCAGGGGCCACCGCCACCATCCCATTTACTATGT 2507

QY 934 lMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAl 954
 Db 2508 TCITGGGATGACAAACGGTTTCACAGAGATGAGCTCCAGATCCTGACGTACCGAGCTGTG 2567
 QY 954 aPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHi 974
 Db 2568 CCACACTTACGTACGATGCACACAGCTCTGTCTCTATCCAGACCTCGCTACTATGCCG 2627
 QY 974 sLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAs 994
 Db 2628 CCTGGTGGCTTTCGGGACGGA---TACCACCTGGTGAGCAAGGAGCATGACAGTGGAGA 2684
 QY 994 p 994
 Db 2685 G 2685

RESULT 3

US-09-533-029-105

; Sequence 105, Application US/09533029

; Patent No. 666446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddie, James

; APPLICANT: Pineda, Omaisra

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 105

; LENGTH: 2910

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G1149

US-09-533-029-105

Alignment Scores:

Pred. No.:	3,39e-56	Length:	2910
Score:	575.50	Matches:	237
Percent Similarity:	39.0%	Conservative:	159
Best Local Similarity:	23.3%	Mismatches:	337
Query Match:	10.8%	Indels:	283
DB:	3	Gaps:	41

US-10-645-746-3 (1-1020) x US-09-533-029-105 (1-2910)

QY 38 LysLysValLeuLeuLeuValAlaValPheLysPheSerSerLysLysLysTyrAspArgGlu 57
 Db 472 AGAAGATCATGGTTCGTGGAAAT-----CATTTCTTGTTCGAAGTTCGTGATCGTAT 525
 QY 58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro 77
 Db 526 CTCACCAATACGATGTTTCGATCAATCCTCGAGTTATATCAAG-----570
 QY 78 PheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
 Db 570 -----570
 QY 98 ArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp 117

Db 570 -----570
 QY 118 ThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLys 137
 Db 571 -----ACAGTGAACAGAAACGTGATGAACCTTCGTGTTAAGAAAT---609
 QY 138 ValValLysLysAspSerGlu-----LysLysAspGluLysAspLeuGluLysLys 154
 Db 610 -----TATAAAGACTCTCACTTGGGAGGGAAGTCACAGCGTATGAT---GGAAGGAA 660
 QY 155 IleLeuTyrThrMet-----IleLeuThrTyrArgLysLysPheHisLeuAsnPheSer 172
 Db 661 AGCCTCTATACTGCTGCTGCTCATTACCTTTTGACCTCGAAAGAGTTGTGTGAATCTGGCG 720
 QY 173 ArgGluAsnProGluLysAspGluGluAlaAsnArgSerTyrLysPhe---LeuLysAsn 191
 Db 721 GAGAAAGAGCTGACGGTTCCTCTGGGAAGGACAGACCGTTTAAAGTTGCTGTGAAGAAAT 780
 QY 192 Val-----MetThrGlnLysValArgTyrAla 200
 Db 781 GTGACAAGCACTGATCTTTATCAGTTGCCAACAGTTCCTTCGATCGTAAGCAAGAGAGGCT 840
 QY 201 PropheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTyrAspAsnAsn 220
 Db 841 CCATAT-----GATACCTATCCAAGTCTTGATGTTGTTCTTAGGGATAAGCCCTCTAAT 894
 QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeu--- 239
 Db 895 GATTATGCTCTGTTGGAGGTCCTTT-----TTCACACTAGTTTGGGA 939
 QY 240 -----GluValAlaProArgIleGluAlaTrpPheGlyLe 251
 Db 940 AAGGACGCAAGAGATGTTAGGGGTGAGCTTGGAGATGGTATTGAGTACTCGAGAGGTTAT 999
 QY 252 TyrIleGlyLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAsp 271
 Db 1000 TTCCAAAGTCTAAGG---CTGACTCAGATGGGTTTGTCTCGAACATTCAGCTTTCAGCA 1056
 QY 272 LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuValAsp 291
 Db 1057 AGATCATTTTAT-----GAA 1071
 QY 292 ProGlnSerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
 Db 1072 CCGATTGTTGTCACCTGACTTTATTAGCAAGTTCCTGAATATAAGGGACTTA-----AAC 1125
 QY 312 MetThrIleArgGlnAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeu 331
 Db 1126 AGGCCACTTAGAGACTCAGATCGACTTAAAGGTGAAGAAAGTTTGTGAGGACACTGAAAGTT 1185
 QY 332 LysCysAlaGluValTrpAspAsnGluMetSerArg-----343
 Db 1186 AGTTGCTTCAC---TGGAAACGCGCACAAAAGTGCCTGCAAAATAGTGGGATTTCTAGTCTA 1242
 QY 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyr 362
 Db 1243 CCCATCAGGAGCTAAGGTTCACTCTGGAGGACAAATCAGAGAAGACGGTTGTTCAATAT 1302
 QY 363 LysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPhe 382
 Db 1303 -----TTTGTGAAAAATATAAT-----1320
 QY 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
 Db 1321 -----TATAGA-----GTGAATACCGAGCTCTACTGCTATTCAACA 1359
 QY 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGlyLysPro 422
 Db 1360 GCGAGTGACAAAGACCCGCTACCTACCAATGGAGCTCTGCCAAATTCAGGAA---GGG 1416
 QY 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442

```
Db      1417 CAAGATACACCAAGGCTCAATGAGAACGAAGTCACTGCATTCGCTAAAGCTACCTGC 1476
Qy      443 ArgLysProHisAspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAaspPheSer 462
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1477 CAACGACCCCTGAT---AGAGAGAACTCGATCAAAACTTGGTTGTGAAAAATAATTAC 1533
Qy      463 SerGluGluLeuAenPheValGluArgPheGlyLeuLysSerLeuLysGluMetIleGlu 482
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1534 AATGATGATCTGAGC-----AAGGAGTTTGGGATGTCAGTCACTACCCCACTAGCCTCG 1587
Qy      483 CysProGlyLysValLeuLysGluProMetLeuValAenSerValAenGluGlnIleLys 502
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1588 ATTGAAGCTCGTGACTTCCCCACCGATGTTGAAGTACCATGACAGTGGTAAAGGAAA 1647
Qy      503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAenValProGluLysGlu 522
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1648 ATGGTAAATCCAAGGCTAGGA-----CAGTGAACATGATTGACAAAGAAA--- 1692
Qy      523 LeuCysCysAlaValPheValValAenGluThrAlaGlyAenProCysLeuGluGluAen 542
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1692 ----- 1692
Qy      543 AspValValLysPheTyrThrGluLeuIleGlyCysLysPheArgGlyIleArgIle 562
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1693 -----ATGGTTAATGGAGCAAAA-----GTCACCTTCT 1719
Qy      563 GlyAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAenGluTyr 582
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1720 TGGACTTGGCAATTTAAGCCTCAACCTGCTATTTCGGTTTCATCTTGTCCCTCGAACAT 1779
Qy      583 AlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaAlaThr 602
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1780 -----ATTGAGAAAGCTCTTCTC 1797
Qy      603 GluAlaLysAenMetPheGluArgLeuProAaspLysGluGlnLysValLeuMetPheIle 622
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1798 GAT-----ATCCACAAAGGGCACCTGGTCTCAA-----CTGTGATT 1836
Qy      623 IleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHisThr 642
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1837 GTAATATTCCCTGATGTCATGCATCATATGGAATAATAAAGGATCTGTGAAACAGAA 1896
Qy      643 IleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeu 662
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1897 TTGGGATTTGCTCTCAGTGTGCCAACCTAGACAAAGTTAATAAATC----- 1944
Qy      663 ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAenAlaLys 682
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1945 -----AACAAGCAGTACATGGAATAATGTTGCCTTGAAGATCAATGTCAAG 1989
Qy      683 LeuGlyGlyIleAenGlnGluLeuAaspTrpSerGluIleAlaGluIleSerProGluGlu 702
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1990 ACTGGGGGAAGGAACACTGTTCTAAT----- 2016
Qy      703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2017 GATGCTATTAGAGAAACAACTCTTATTACTGATCGTCCAACCAACATCATCGGTGCT 2076
Qy      718 AspValThrHisProThrSerTyrSerGlyIleAepTyrSerIleAlaAlaValValAla 737
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2077 GATGTGACTCACCCACAGCTGGAGAGGACTCAAGTCCCTTATTGCTGCTGTGTGGCC 2136
Qy      738 SerIleAen---ProGlyGlyThrIleTyrArgAenMetIleValThrGln----- 753
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2137 TCTATGGAGTGGCTCGAGATAACAAATAACCGAGGATTTGGTTCTGCTCAAGCTCATAGG 2196
Qy      754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2197 GAAGAAATTTATTCAGGACCTGTATTAAGCTGGTTCCAGATCCACAGCTGGGTAGTCCAC 2256
Qy      766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2257 -----TCTGGTTTGTAGAGGAACATTTTCATAGCATTCAGGAGAGCT----- 2298
```

```
Qy      786 GluAenAenAaspAenArgAlaProAlaHisIleValValTyrArgAspGlyValSerAasp 805
Db      2299 -----ACAGGCCAGATACCTCCANAGGATCATCTTCTATCGTGACGAGTAAGCGAA 2349
Qy      806 SerGluMetLeuAeqValSerHisAaspGluLeuAeqSerLeuLysSerGluValLysGln 825
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2350 GGGCAGTTTAGTCAGGTTCTGCTACATGAGATGCTATCCGCAAGCCTTGTAACCTCT 2409
Qy      826 PheMetSerGluArgAaspGlyGluAaspProGluProLysTyrThrPheIleValIleGln 845
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2410 CTCCAA-----GAGAAATTATGTTCTCGTGTTACTTTCGTGATTGTCAG 2454
Qy      846 LysArgHisAenThrArgLeuLeuArgMetGluLysAaspLysAaspValValAenLys 865
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2455 AAACGTCACACACACAGTTTGTTCCTCAGCAACACGGG-----AATCGT 2499
Qy      866 AspLeuThrProAlaGluThrAaspValAlaValAlaAlaValLysGlnTrpGluGluAasp 885
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2500 GATATGACT----- 2508
Qy      886 MetLysGluSerLysGluThrGlyIleValAenProSerSerGlyThrThrValAaspLys 905
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2509 -----GATAAGAGTGGCAATATTCAACCA-----GGTACTGCTCGGACACT 2550
Qy      906 LeuIleValSerLysTyrLysPheAaspPhePheLeuAlaSerHisGlyValLeuGly 925
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2551 AAAATCTCTCACCTTAATGAATTCGACTTCATTATTTGAACAGCCATGCTGGTATT 2610
Qy      926 ThrSerArgProGlyHisTyrThrValMetTyrAaspAaspLysGlyMetSerGlnAaspGlu 945
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2611 ACNAGCAGCCGGCACATTACCATGTACTTCTCGATGAGACGGTTTCCACCGCTGATCAG 2670
Qy      946 ValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSer 965
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2671 TTGCAATAGCTCACAAACACCTCTGTACACGTATGCGAGGTGTACAAAATCTGTGTCA 2730
Qy      966 LeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArg 985
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2731 ATTTGGCCACCAACCGCTACTAGCTCATTGGCTGGCTCCGTCGCC-----CQC 2778
Qy      986 ThrTyrLysGluHisTyrIleGlyAaspTyrAlaGlnProArgThrArg 1001
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2779 TACTACATGGAGAGTGAGATGCTCATGGAGGTTCCGACGAGTCCAGG 2826

RESULT 4
US-09-270-767-11928
; Sequence 11928, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11928
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11928

Alignment Scores:
Pred. No.:      3,348-42      Length:      1662
Score:          452.50      Matches:    156
Percent Similarity: 43.8%      Conservative: 91
Best Local Similarity: 27.7%      Mismatches: 213
Query Match:      8.5%      Indels:     104
DB:              3          Gaps:       18

US-10-645-746-3 (1-1020) x US-09-270-767-11928 (1-1662)
```

QY 467 AsnPheValGluArgPheGlyLeuCySerSerLysLeuGlnMetIleGluCysProGlyLys 486
 Db 22 TCGATGTCGAGAGTTGGCTCCATTCCTCAATTCGATGATGAGGATACGAGGACGC 81
 QY 487 ValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal 506
 Db 82 GTCTTGCTCTCTCCAGCTTCAGTATGGGGACGT-----GTG 120
 QY 507 IleArgPheGlnGluLysGlnLeuAsnValValProGluLysGluLeuCysAla 526
 Db 121 TCTACCGGCTCACCGCCAGCAGCTG---TTCGCCGCACAGAACAGGTGAGCTTGCC 177
 QY 527 ValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAspValValLys 546
 Db 178 -----TCCCCACACAGGGTGTGGGATATGCGCGGCAACGAG 216
 QY 547 PheTyrThrGluLeuIleGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGlu 566
 Db 217 TTCTTCACTGGCGTC-----GAGATCCGATCTGGGCCATCGCC 255
 QY 567 AsnArgGlyAlaGln---SerIleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyr 585
 Db 256 TGTGTTTCCCCACAGCGCACGGTGGCGGAGGATGCGCTGGGTAAATTTCAACCAGCAGCTG 315
 QY 586 LysAsnCySerThrLeuAsnThrGly-----IleGly-----ArgPhe 597
 Db 316 CAGAGATCTCAACGATGAGGATGCGGATGATGGACAGCGTGCTTCTGTAAGTAC 375
 QY 598 GluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLys 617
 Db 376 GCCACCGGCGGATCAAGTGGGAACCATGTTCCGTTACCTGAAGTACATCTCCCGCGC 435
 QY 618 ValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis 637
 Db 436 CTGCAACTCGCTCGTGTGTGCTGCCCGGCAAGACTCCAGTATACGCCGAGGTGAAGCGT 495
 QY 638 TyrCyAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
 Db 496 GTAGGTGACACCGTTCTGGGTATGGCCACCATGTTGTGAGCGGCCAAGACGTTGAACAAG 555
 QY 658 AlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeu 677
 Db 556 ACA-----TCCACACAGAGCGCTCTTAATCTGTGTCTG 588
 QY 678 LysIleAsnAlaLysLeuGlyIleAsnGlnLeuLeuAspTrpSerGluIleAlaGlu 697
 Db 589 AAGATCAAGCTCAAGTTGGCGGCATCAATTCAATCTG----- 627
 QY 698 IleSerProGluLysGluArgGlySerThrMetProLeuThrMetTyrValGlyIle 717
 Db 628 ---GTCCCTTCCACTCGCCCAAAGGCTTCAATGAGCCGGTC---ATCTTTTGGGTGCC 681
 QY 718 AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValAla 737
 Db 682 GATGTGACACACCCACCGCTGGCGACACAAAGAACCATCGATTGCGCGCTGCTGGGC 741
 QY 738 SerIleAsnProGlyThrIleTyrArgAsnMetIleValThrGlnGluCysArg 757
 Db 742 TCCATGGATGCCCATCCATCGCGGTATGCGCCACCGTTCGCGTACAGCAG----- 792
 QY 758 ProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPhe 777
 Db 793 -----CACCACACAGCATCATCCAGGAGCTGAGCAGCATGGTG 831
 QY 778 ValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAlaHisIleVal 797
 Db 832 CGCGACTGTTGATCATGTTCTACAGTGCAGCGGGGTACAGGCCCAACCCCATCATATA 891
 QY 798 ValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArg 817
 Db 892 CTCATCGTACGAGGAGTCTCCGAGGAGCAATTCACATGTCCTGCAACACGAAATTGACC 951
 QY 818 SerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGlu---- 836

Db 952 GCCATTCCGAGGCGCTCATTAAGCTA-----GAGCCAGAGTAT 990
 QY 837 ---ProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgArg 855
 Db 991 CGGCGGGGCATCATTTTCATTTGTTGTCAGAGAGGCCATCATCTCGACTCTTCTGCGCG 1050
 QY 856 MetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAla 875
 Db 1051 GAGAAGAAGGAG----- 1062
 QY 876 ValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleVal 895
 Db 1063 -----CAGAGCGCAATCGGC----- 1080
 QY 896 AsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhe 915
 Db 1081 AATATTTCCCGCAGGACCATCCGTCGATGTGGGCATCACATCCACCGAATTTGATTTTC 1140
 QY 916 PheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMet 935
 Db 1141 TATCTCTGCAGCCATCAGGGCATCCAGGCACCATGTCGCCCTCGCACTACCACTGCTCTG 1200
 QY 936 TyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPhe 955
 Db 1201 TGGGACGACAATCACTTTTGACTCGGACGAGCTGCAGTGCCTCATGCTATGCTATGCTAT 1260
 QY 956 LeuSerAlaAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHisLeu 975
 Db 1261 ACGTACGTGCTGTCACCGCATCGTCAGTATACCGGCGGCAGCTACTACGCCCATTTA 1320
 QY 976 SerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAsp-Ty 995
 Db 1321 GTGCGCTTCGTGCCAG---TATCATCTGTCGAGAAGGAGCAC-----GATTTCG 1368
 QY 995 rAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThrAsnValLysTyrPr 1015
 Db 1369 GCGGAGGGTTCGACCAAGAGCGGCTGCTCAGAGGATGCTAGC---CAGGTGCCATGGCC 1425
 QY 1015 oGlyMetSer 1018
 Db 1426 AGGGCCATCA 1435
 RESULT 5
 US-09-270-767-12369
 ; Sequence 12369, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12369
 ; LENGTH: 1874
 ; TYPE: DNA
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-12369
 Alignment Scores:
 Pred. No.: 1,5e-34 Length: 1874
 Score: 388.00 Matches: 156
 Percent Similarity: 38.5% Conservative: 107
 Best Local Similarity: 22.8% Mismatches: 269
 Query Match: 7.3% Indels: 151
 DB: 3 Gaps: 23
 US-10-645-746-3 (1-1020) x US-09-270-767-12369 (1-1874)
 QY 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAsp----- 378

Db 182 GTATACCGTGAATGGTCTTTCTCGTGGCCCGCCAGCAGTGACACCTTCGAAACACGAT 241
 Qy 379 -----ThrThrLeuPheLysIleTyrGluGluAsnLysPheIleGluPhePro 395
 Db 242 GGGAGAAAGGTCACTTCTTCTTCCACAGTCGCAACTACCCATTGAAGTTTCC 301
 Qy 396 HisLeuProLeuValLysValLysSerGlyAlaLysGluTyrAlaValProMetGluHis 415
 Db 302 CAACTCCATTGTCTGAACCTTGGATCTTCAATCAAGAGTATTCTGTGCCCATCAGGTA 361
 Qy 416 LeuGluValHisGluLysProGlnArgTyrLysAsnArgIleAspLeuValMetGln--- 434
 Db 362 TCGAGCATCGAGAA-----GGTCAGGCCCTAAACCGCAAGGATCGAGCAACTCAGGTG 415
 Qy 435 AspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAsnThrLeuLys 454
 Db 416 GCCAATATGATAAAGTACGACGCCACATCGACGAACGTCGGAAGCGCAAGATTATGAAC 475
 Qy 455 MetLeuLysGluLeuAspPhe-SerSerGluGluLeuAsnPheValGluArgPheGlyLe 474
 Db 476 TTGCTGCAATAC-----TTCCAGCACAAACCTGGGATCCGACCATCAGTCGCTTTGGCAT 529
 Qy 474 uCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeuVa 494
 Db 530 CCGCATTCGCAACGATTTTATTTGGTGAAGCACCCCGCTCTTAACCCACCTCAGGTT-- 587
 Qy 494 IAsnSerValAsnGluGlnIleLysMetThrProValIleArgLysPheGlnGlu----- 512
 Db 588 -----GAATATCATAGTAAGAGGTTTACTATGTGTGAAGACGGGTCTGTGGCGCATGGA 640
 Qy 513 -----LysGlnLeuAsnValProGluLysGluLeuCysCysAlaValPheVal-- 529
 Db 641 TGGCATGAAGTTTCTGGAGCCCAAGCCCAAGCGGCACAGTGTGGCTTGTATTTGCGA 700
 Qy 530 -----ValAsnGluThrAlaGlyAsnProCy 538
 Db 701 TCCGAGGAGTGTGCGAAATGAACCTATGCCCCAGCTGAATGAC---TTCCGGGAACCTAAT 757
 Qy 538 sLeuGluGluAsnAspValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheAr 558
 Db 758 AATATCCCAAGGCAAGGCGTCAACATA-----ACCTTGGATTCTGTGACATACAG 811
 Qy 558 gGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaTh 578
 Db 812 ACCGTT-----ACGGATGCGAACCGACGCTAGACATATTTTCGCGATCTGAA 862
 Qy 578 rLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGl 598
 Db 863 GCGCAGCCAGCAGAT----- 878
 Qy 598 uIleAlaIleThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysVa 618
 Db 878 ----- 878
 Qy 618 lLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTy 638
 Db 879 -CTGCAATTGTGATTATCTCTCAGTTAGATT---TCCTACGATACAAATTAAACAGAA 934
 Qy 638 rCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAl 658
 Db 935 GCGCAGCTGCGACGATGGAATTTTGAAGCAATGCAATTAAGCAGTTTCAACGCTGGAACGA-- 992
 Qy 658 lLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLys 678
 Db 993 -----AAGTGTAAATATCAGACGATTGGAAATATTTCTACTTAA 1030
 Qy 678 sIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTyrPheGluIleAlaGluI 698
 Db 1031 GATCAACTCCAGCTGAAGCGGATCAACCAAGATC----- 1067
 Qy 698 eSerProGluGluLysGluArgLysThrMetProLeuThrMetTyrValGlyIleAs 718
 Db 1068 -----AAGGATTCCTCGTCTGCGGATGATGAAGAACACCATGATGATGGAGCGCA 1120

Qy 718 pValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlase 738
 Db 1121 TGTGACCATCCCTCTCCGATCAAGCGAGATTCCTGATGGTGGGATGACAGCCTC 1180
 Qy 738 rIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysArgPr 758
 Db 1181 ACAGCATCCCTACGAGCGCAGTTAATCAATGCAATATCGTTTGCAG----- 1226
 Qy 758 oGlyGluArgAlaValAlaHisGlyArgLysGluThrAspIleLeuGluAlaLysPhe-- 777
 Db 1227 -----CGAGGGCTCTCGAGGAGATTGAGACATGTTCTCTC 1261
 Qy 778 -----ValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAlaHi 795
 Db 1262 GATTACTTTGGAGCACTTGGCGGTGTATAAGAGATACCGT---AACGCCCTATCTTGATCA 1318
 Qy 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGl 815
 Db 1319 TATCATCTACTACCGAGATGGGTGAGCGCGCCAGTTTCCGAAAAATCAAAAACGAGGA 1378
 Qy 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
 Db 1379 ACTGAGGTGTATTAACAACAGCCTGT-----GACAAGGTGGGCTG 1417
 Qy 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgAr 855
 Db 1418 TAAACCCAGAGATTTGTCGTGATTTGGTGAAGCGTCATCACCTCTTCTTCCC--- 1475
 Qy 855 gMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAl 875
 Db 1475 ----- 1475
 Qy 875 aValAlaValLysGlnTrpGluAspMetLysGluSerLysGluThrGlyIleVa 895
 Db 1476 -----AGCGGCGACGTAACGACATCGAAACAAGTTTCAACAACGT 1513
 Qy 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
 Db 1514 GGACCCC-----GGAACCGTGTGATCGACACCATTTGTCATCTCAACGATGATGAT 1567
 Qy 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
 Db 1568 CTTTCATGTGACGCCACCGCCATCCAGGCGACGCGCAAGCCCAACACGATACAAATGTGAT 1627
 Qy 935 tTyrAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955
 Db 1628 TGAGAACACAGGCAATCTTGACATCGACTTGTTCAGCAGTTTGACCTACAACCTGTGCCA 1687
 Qy 955 eLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHisLe 975
 Db 1688 CATGTTCTCTGTTGCAATCGCTCGGTTTCTTATCGGCTCGGCTTATTGACCATTT 1747
 Qy 975 uSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGly----- 993
 Db 1748 GGTAGCTGCTCGTGGACGC---GTTTATCTGATGCGCAACAGGTCTCTGATTTGAA 1804
 Qy 994 -----AspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThrAsnVa 1012
 Db 1805 GAAGGAGTACGCAAG-----CGAACGATTTGTCGCGAATTCATGAAGAAAAACCC 1855
 Qy 1012 lLysTyr 1014
 Db 1856 CATGTAC 1862

RESULT 6

US-10-043-774B-1
 ; Sequence 1, Application US/10043774B
 ; Patent No. 690017
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Illinois at Chicago
 ; APPLICANT: Sharma, Arun
 ; APPLICANT: Hoffman, Ronald

; TITLE OF INVENTION: HUMAN HEMATOPOIETIC GROWTH REGULATORY GENE AND USES

; FILE REFERENCE: MBHB: CU08/PPA
 ; CURRENT APPLICATION NUMBER: US/10/043,774B
 ; CURRENT FILING DATE: 2002-06-06
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 2328
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2328)
 ; OTHER INFORMATION: Human Hiwi Protein
 US-10-043-774B-1

Alignment Scores:

Pred. No.: 1,78e-25 Length: 2328
 Score: 312.00 Matches: 150
 Percent Similarity: 36.7% Conservative: 88
 Best Local Similarity: 23.1% Mismatches: 218
 Query Match: 5.8% Indels: 192
 DB: 3 Gaps: 27

US-10-645-746-3 (1-1020) x US-10-043-774B-1 (1-2328)

Qy 417 GluValHisGlu---LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAsp 435
 Db 613 GAAGAATAAATTTCAAGAACAAAGTTTCCAAAGAACTAATAGTTTGTCTTACC 672
 Qy 436 LysPheLeuLysArgAlaThrArgLysProHis---AspTyrLysGluAsnThrLeuLys 454
 Db 673 AAGTATAACATTAAGACATACAGATGGATGATTTGCTGGGACCAAGAAATCCCAAGAGC 732
 Qy 455 MetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheValGluArgPheGlyLeu 474
 Db 733 ACCTTTAAGAAAGCCGAGGCTCT-----GAAGTCAGCTTCTTAGAATACTACAGGAAG 786
 Qy 475 CysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeuVal 494
 Db 787 CAATACAACCAAGAGATACCCGAC-----TTGAAGCAGCCTGTCTTGGTC 831
 Qy 495 AsnSerValAsnGluGln-----ile 501
 Db 832 AGGCAGCCCAAGAGAGGGGGGCTTGGGGGACACTGCCAGGGCTGCCATGTCTCATT 891
 Qy 502 LysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsn-----516
 Db 892 CCTGAGCTCTGCTATCTTACAGGTCTAACTGATAAAATGCGTAATGATTTTAACGTGATG 951
 Qy 517 -----ValValProGluLysGluLeuCysCysAlaValPhe 528
 Db 952 AAAGACTTACCGCTTCCATCAAGACTAACTCCAGAGCAAAAG-----993
 Qy 529 ValValAsnGluThrAlaGlyAsn-----ProCysLeuGluGluAsnAspValValLys 546
 Db 994 -----CAGCGTGAAGTGGACGACTCATTTGATTACATTAATAAAACCAATATGTTCAA 1047
 Qy 546 -----546
 Db 1048 AGGGAGCTTCGAGACTGGGGTTTGAGCTTTGATTCCACTTACTGTCTCTCAGGAAGA 1107
 Qy 547 ----PheTyrThrGluLeuIle-----GlyGlyCysLysPhe-----ArgGlyIle 560
 Db 1108 ATTTTGCACAAACAGAAAGATTTTCAACCAAGGTGGAAAAACATTTGATTACAATCCACAATTT 1167
 Qy 561 ArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsn 580
 Db 1168 GCAGATTGGTCCAAAGAAACAGAGGTGCACCAATTAATTAATAGTTTAAAGCCACTTAGATAAC 1227
 Qy 581 GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAla 600
 Db 1228 TGGCTGTTGATCTATACGCGA-----AGAAATTAATGAGCAGGCC 1266

Qy 601 AlaThrGluAlaLysAsnMetPheGluArgLeuPro-----612
 Db 1267 AATTATTGATACAAAATCTATTTAAAGTTTACACGACCATGGCGCATGCAAAATGAGAAA 1326
 Qy 613 -----AspLys-----GluGlnLys 617
 Db 1327 GCATATATGATTGAAGTGGATGACAGACGTAAAGCCCTACTTAAGAGTCTTACAGCAAAAG 1386
 Qy 618 Val-----LeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyr 632
 Db 1387 GTCACAGCAGACACCCAGATAGTTGTCTGTCTGTGTCAAGTAATCGAAGCAAAATAC 1446
 Qy 633 GlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSer 652
 Db 1447 GATGCTATTAAAAAATACCTGTGTACAGATTGCCCTACCCCAAGTCAAGTGTGTGTGGCC 1506
 Qy 653 GluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePhe 672
 Db 1507 CGAACCTTAGGCAAA-----CAGCAAACTGTCATGGCCATTTGCT 1545
 Qy 673 TyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTyr 692
 Db 1546 ACAAGATTGCCCTACAGATGAAGTCCAGATGGGGA-----GAGCTC---TGG 1593
 Qy 693 SerGluIleAlaGluIleSerProGluGluLysGluArgArgLysThrMetPro-----710
 Db 1594 -----AGGTGGACATCCCTCCTGAAG 1614
 Qy 711 LeuThrMetTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyr 730
 Db 1615 CTGCTGATGATCTTGGCATGATGTTACCATGACATGACAGCTGGG-----CGAGG 1668
 Qy 731 SerIleAlaAlaValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIle 750
 Db 1669 TCAATCGCAGGATTTGTTGCCACATCAATGAAGGATGACCCGCTGTTCTCAGCTGC 1728
 Qy 751 ValThrGlnGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThr 770
 Db 1729 ATATTTTCAGAT-----AGAGCAGCAGGAGCTGCTA 1758
 Qy 771 AspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAspAsn 790
 Db 1759 GATGGGCTCAAGCTCCCTCGCAAGCGGCTCTCAGGGCTTGGAAATAGCTGCAATGAG---1815
 Qy 791 ArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArg 810
 Db 1816 TACATGCCAGCCGCGATCATCGTGTACCGCGATGGCGTAGGACACGCCAGCTGAAACA 1875
 Qy 811 ValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSer-----828
 Db 1876 CTG-----GTGAACACGAAGTGCACAGTTTTTTGGATTGTCTA 1914
 Qy 829 GluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHis 848
 Db 1915 AAATCCATTGGTAGAGTTTACACCCCTAGACTAACCGTAAATTTGTTGGTGAAGAAAGAGT 1974
 Qy 849 AsnThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThr 868
 Db 1975 AACACCAAGATTTTT-----1989
 Qy 869 ProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLysGlu 888
 Db 1989 -----1989
 Qy 889 SerLysGluThrGly-----IleValAsnProSerSerGlyThrThrValAspLysLeu 906
 Db 1990 ---GCTCAGTGGGAGGAAGACTTCAGAATCCACTTCTCGGAACAGTTATTGATGAGAG 2046
 Qy 907 IleValSerLysTyrLysPheAspPheLeuAlaSerHisGlyValLeuGlyThr 926
 Db 2047 GTTACCAGACCAAGATGTTGATCTTTTTTATCGTACGCCAGCTGTGAGAAAGTGTAGT 2106

QY 927 SerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluVal 946
Db 2107 GTTCTCCACACATTACAAATGTCATCTATGACACAGCGGCTGAAGCCAGACACATA 2166
QY 947 TyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeu 966
Db 2167 CAGCGCTGACCTACCAAGCTGTGCCACATCTATTACAACTGCCAGGTGTCATTCTGT 2226
QY 967 ProValProValHisTyrAlaHis 974
Db 2227 CCTGCTCCTTGCAGTACGCCAC 2250

RESULT 7

US-09-873-737A-5
; Sequence 5, Application US/09873737A
; Patent No. 6723534
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE
; FILE REFERENCE: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/873,737A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110,901
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (167)..(2749)
; NAME/KEY: misc_feature
; LOCATION: (392)
; OTHER INFORMATION: n=t or a, Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (1073)
; OTHER INFORMATION: n=c or a, Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (2369)
; OTHER INFORMATION: n=c or a, Xaa=Leu or Ile
US-09-873-737A-5

Alignment Scores:
Pred. No.: 1,81e-24 Length: 3472
Score: 306.00 Matches: 151
Percent Similarity: 37.4% Conservative: 88
Best Local Similarity: 23.6% Mismatches: 226
Query Match: 5.7% Indels: 174
DB: 3 Gaps: 26

US-10-645-746-3 (1-1020) x US-09-873-737A-5 (1-3472)

QY 417 GluValHisGlu---LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAsp 435
Db 1037 GAAGACATAAATTTCAAGAACACAGTTTCCAAAGAANTATAGTTTAGTCTCTTACC 1096
QY 436 LysPhe---LeuLysArgAlaThrArgLysProHisAspTyrLysGluAsnThrLeuLys 454
Db 1097 AAGTATAACATTAAAGACATACAGAGTGGATGATATTGACTGGGACCAAGATCCCAAGAGC 1156
QY 455 MetLeuLysGluLeuAspPheSer----- 462
Db 1157 ACCTTTAAGAACCCGACCGGCTCTGGGGTCACGCTCTTAGAATACTACAGGAAGCAATAC 1216
QY 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu 482
Db 1217 AACCAAGAGATCACCAGCTTGAAGCAGCCCTGCTTGGTCAGCCAGCCCAAGAGAGCGG 1276

QY 483 CysProGlyLysValLeuLysGluPro---MetLeuValAsnSerValAsnGluGlnIle 501
Db 1277 GGCCTTGGGGGACATCGCAGGCTCGCATGCTCATTTCTGAGCTGCTATCTTACA 1336
QY 502 LysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLys 521
Db 1337 GGTCTAACTGATAAATGCGT-----AATGATTTTAACGTGATG-----AAA 1378
QY 522 GluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCys 538
Db 1379 GACTTAGCGGTTTCATACAGACTAATCCAGAGCAAGGCGAGCTGCGAGCTGGGTTTGAGC 1438
QY 539 -----LeuGluGluAsnAspValValLys----- 546
Db 1439 ATTGATTACATTTATAAAACGATAATGTTCAAAGGGAGCTTCGAGACTGGGTTTGAGC 1498
QY 547 -----PheTyrThrGluLeuIle----- 552
Db 1499 TTTGATTCCAACTTACTGCTCTCTCAGGAAGAAATTTGCCAAACAGAAAAGATTACACAA 1558
QY 553 GlyGlyCysLysPhe-----ArgGlyIleArgIleGlyAlaAsnGluAsnArgGly 569
Db 1559 GGTGGAAGAAACATTTGATTACAATCCACAATTTGCAGATTGGTCCAAAGAAACAGAGGT 1618
QY 570 AlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThr 589
Db 1619 GCACCATTAATTAGTGTAAAGCCACTAGATAACTGGCTGTGATCTATATACGCGA----- 1672
QY 590 LeuAsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGlu 609
Db 1673 -----AGAAATATGAGCAGCCCAATTCATTGATACAAAATCTATTAAA 1717
QY 610 ArgLeuPro----- 614
Db 1718 GTTACACAGCCATGGGCATGCAAAATGAGAAAAGCAATAATGATTGAAGTGGATGACAGA 1777
QY 615 -----GluGlnLysVal-----LeuMetPhe 621
Db 1778 ACTGAAGCCTACTTAAGAGTCTTACAGCAAAAGGTCACAGCAGACACCAGATAGTTGTC 1837
QY 622 IleIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAspHis 641
Db 1838 TGTCTGTTGTCAGTAATCGGNAGGCAAAATACGATGCTATTAAAAAATACCTGTGTACA 1897
QY 642 ThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSer 661
Db 1898 GATTGCCCTACCCCAAGTCAGTGTGTGTGGGCCCGAACCTTAGGCAAA----- 1945
QY 662 LeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAla 681
Db 1946 -----CAGCAAACTGTATGCCCATTTGCTACAAAGATTGCCCTACAGATGAAGTGC 1996
QY 682 LysLeuGlyGlyIleAsnGlnGluLeuAspTyrSerGluIleAlaGluIleSerProGlu 701
Db 1997 AAGATGGGAGGA-----GAGGTC-----TGG----- 2017
QY 702 GluLysGluArgArgLysThrMetPro-----LeuThrMetTyrValGlyIleAspVal 719
Db 2018 -----AGGGTGGACATCCCTCGAAGCTCGTGATGATCGTTGGCATCGATTGT 2065
QY 720 ThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlaSerIle 739
Db 2066 TACCATGATGATGACCGCTGGTTCTCAGCTGCATATTTTCAGAT----- 2119
QY 740 AsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysArgProGly 759
Db 2120 AATGAAGGATGATCCCGCTGGTTCTCAGCTGCATATTTTCAGAT----- 2164
QY 760 GluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLys 779
Db 2165 -----AGAGGACAGGAGCTGGTAGATGGGCTCAAAAGTCTGCTGCAAGCG 2209

```

QY 780 LeuLeuArgGluPheAlaGluAsnAspAsnArgAlaProAlaHisIleValValTyr 799
Db 2210 GCTCTGAGGGTTGGAAATAGCTGCAATGAG---TACATGCCACCGCGGATCATGCTGATC 2266
QY 800 ArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeu 819
Db 2267 CGCGATGGCGTAGGAGACGGCCAGCTGAAACACATG-----GTG 2305
QY 820 LysSerGluValLysGlnPheMetSer-----GluArgAspGlyGluAspProGluPro 837
Db 2306 AACTACGAAGTGGCCACACATTTTGGATGTTCTAAATATCCATGCTGAGAGGTTACAACTT 2365
QY 838 LysTyrThrPheIleValIleGlnLysArgHisenThrArgLeuLeuArgArgMetGlu 857
Db 2366 AGANTACGGTAATGTGGTGAAGAAAGAGTGAACACCACTTTT----- 2413
QY 858 LysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAla 877
Db 2413 ----- 2413
QY 878 AlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGly-----IleVal 895
Db 2414 -----GCTCAGTCTGGAGGAAGACTTCAG 2437
QY 896 AsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhe 915
Db 2438 AATCCACTTCTGGAACAGTTATGATGTAGAGTTTACAGACAGACAGAAATGGTATGACTTT 2497
QY 916 PheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMet 935
Db 2498 TTTATCTGAGCCAGCGCTGTGAGAGTGTAGTGTCTTCTCCACACATTACATGATCATC 2557
QY 936 TyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPhe 955
Db 2558 TATGACAAACAGCGCGCTGGAAGCCAGACCATACAGCGCTTGACCTACAAGCTGTGCCAC 2617
QY 956 LeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHis 974
Db 2618 ATCTATTACACTGGCCAGGTGTCATTCGTGTTCTCTGCTCTCTGCGCAGTACGCCAC 2674

RESULT 8
US-09-873-737A-3
; Sequence 3, Application US/09873737A
; Patent No. 6723534
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE
; TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6723534 180-104/2
; CURRENT APPLICATION NUMBER: US/09/873,737A
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4064
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (191)..(2776)
; NAME/KEY: misc feature
; LOCATION: (458)
; OTHER INFORMATION: n= a or c, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (836)
; OTHER INFORMATION: n= a or c, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (1337)

```

```

; OTHER INFORMATION: n= a or c, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (2636)
; OTHER INFORMATION: n= a or c, Xaa=Leu or Ile
US-09-873-737A-3

Alignment Scores:
Pred. No.: 4,08e-24 Length: 4064
Score: 304.00 Matches: 151
Percent Similarity: 38.0% Conservative: 109
Best Local Similarity: 22.1% Mismatches: 247
Query Match: 5.7% Indels: 178
DB: 3 Gaps: 26

US-10-645-746-3 (1-1020) x US-09-873-737A-3 (1-4064)

QY 414 GluHisLeuGluValHisGlu---LysProGlnArgTyrLysAsnArgIleAspLeuVal 432
Db 1055 CAGCAGACAGAGGAGCAAGTTCCAGAGCAAGTGTCCAGAGGAGCTCATAGGCTCATC 1114
QY 433 MetGlnAspLysPheLeuLysArgAlaThrArgLysProHis---AspTyrLysGluAsn 451
Db 1115 GTTCTCACCAGTACATAACAAGACCTACCGGTGGATGACATTGACTGGGACCAGAT 1174
QY 452 ThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheValGluArg 471
Db 1175 CCAAGAGCACCTTCAAGAGGCGGATGGCTCG-----GAGGTGAGCTTCTCTGGAGTAC 1228
QY 472 PheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluPro 491
Db 1229 TACAGGAGCAATACAACAGAGGATCACGGAC-----CTGAGCAGCGCG 1273
QY 492 MetLeuValAsnSerValAsnGluGln-----ValValProGlu-----LysGlu 522
Db 1274 GTGCTGTGAGCCACCAAGCGAGGAGGCGCCCGCGGACCCCTGCTGGAGCGGCGGAG 1333
QY 501 -----IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsn----- 516
Db 1334 ATGTCATCTCTGAACTCTGCTATCTCAGAGGCTGACTATAAATGCGCAATGATTTC 1393
QY 517 -----ValValProGlu-----LysGlu 522
Db 1394 AATGTGATGAGACCTGGCAGTGCACACCGGCTGACCCCTGAGCAGCGGCGGAG 1453
QY 523 LeuCysCysAlaValPheValAsnGluThrAlaGlyAsnProCysLeuGlu----- 540
Db 1454 GTGGCGCGCTCATCGACTACATCCACAAGAT---GACATGTGCAGAGAGACTTCGA 1510
QY 541 -----GluAsnAspValValLysPheTyrThrGluLeuIle----- 552
Db 1511 GACTGGGGCTGAGCTTCGACTCAAACTGCTCTCTCTCTGGAAGAAATCTTACAATCT 1570
QY 553 -----GlyGlyCysLysPhe-----ArgGlyIleArgIleGlyVala 564
Db 1571 GAGAAATCCACAGCGCGGGAAGAGCTTTGATTACACCACAAATTGGCAGACTGGTCC 1630
QY 565 AsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyrAlaPhe 584
Db 1631 AAAGAAACAAGAGCGCGCGCTGATCAGCGTGAAGCCATTGGATAACTGCTGCTGATC 1690
QY 585 TyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaThrGluAla 604
Db 1691 TATACCCGC-----AGGAATTATGAGCAGGCCAACTCACTGATGATA 1729
QY 605 LysAsnMetPheGluArgLeuPro-----GluGlnLysVal----- 612
Db 1730 CAGAACCTGTTCAAGTGACTCCAGCGATCGGCATCCAGATGAAAGGCAATCATGATC 1789
QY 613 -----AspLys-----GluGlnLysVal----- 618
Db 1790 GAGGTGATGACAGACAGAGAGCTTATCTGAGAGCCTTTCGAGCAGAGGTCAGCTCAGAC 1849
QY 619 -----LeuMetPheIleIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLys 636

```


1032	DB	GAATCAGCATTAAGATTACAGATGACAAAAG-----ATCTGGAGCCCTGCTCTGACCTG	1085
224	QY	ArgValPro-----	226
1086	DB	TGCATTCCCTTCTCAACATGTTGTGTTTTCGTCGGGTAAATGAACATTTTAGATATGACGCTT	1145
227	QY	-----GluSerPheHisAspProAsnArg-----PheGluGlnSerLeuGluVal	241
1146	DB	GTGGGGAGAAACTTTTATGACCCCTACAAAGTGTCTGTTACTGTACAGCAACAC-----	1196
242	QY	AlaProArgIleGluAlaThrPheGlyIleTyrIleGlyIleLeuLeuPheAspGly	261
1197	DB	-----AGATTGGAGATCTGCCAGGCTATCGAGCTAGCATCTCCGAAGACA-----GATGGA	1247
262	QY	GluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMetSer	281
1248	DB	GGGCTCTTCCTGCTAGCTGATGTCCTCCATAAGGTCAATTCGGATGAC-----TGT	1298
282	QY	LeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAspAspValArgLys	301
1299	DB	GTGCTGGATGTATCATGCATGCCATTTATCAGCAG-----AATAAGAACACTTCACG	1349
302	QY	AspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaIleAspProArg	321
1350	DB	GATGAGTGTACTAAGATCTCTCGTTGGCAATATCTTATCACCCGA-----	1394
322	QY	IleArgGlnLeuLeuGluAsnLeuLysCysAlaGluVal-----TTPAsp-----	338
1395	DB	-----TATAACATCGTACCTATCGTATGATGATGGATTTGGAATAAGACT	1442
339	QY	-----AsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGlu	356
1443	DB	CCAAAGGATAGCTTTCACGATCTCTGATGGGAAGAGATCATCTTCTTGGAACTACACGC	1502
357	QY	GluAsnSerLeuValTyrLysValThrGlyLysSerAspArg-----Gly	371
1503	DB	AAAAAT-----TATGGGATCACAGTTAAGGAAGAGGACCAAGCCATTCGTGATTCAC	1553
372	QY	ArgAsnAlaLysLysTyrAspThr-----ThrLeuPheLysIleTyrGluGluAsnLys	389
1554	DB	AGGCCAGTGAGACACAGGATAATCATGGGATGCTGTAAAA-----	1595
390	QY	LysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyr	409
1596	DB	-----GGGGAAATCCTGCTGCTGCTGCTGAGCTTCTTTTATGACCGCAATCCAGAG--	1646
410	QY	AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIle	429
1647	DB	-----AAGATGAAGAAGGACTTCAGAGCCATGAAG	1676
430	QY	AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys	449
1677	DB	GATTGGCTCAGCAA-----ATCAATCTGAGCCCCAAGCAACCACTAGTCTTTG	1727
450	QY	GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal	469
1728	DB	GAATGCTTGCTGCAAGAAATTGCAAGAACGAGGACGCCACCAATGAACG-----	1778
470	QY	GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys	489
1779	DB	ATGGTTGGGGGCTTCGCTGCAAAAAGGATGTACATAAGATTGAAGACGCTGTTCTG--	1835
490	QY	GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProValIleArgGly	509
1836	DB	-----CCAATG-----GAAAGAAATTAAAC-----TTAAAAAATACCTCGTTTATC-----	1874
510	QY	PheGlnGluLysGlnLeuAsnValValProGluLysGluLeuCys-----	524
1875	DB	-----ACATCTCAGGAACATAACTGGGTT-----AAGGAAGTAACAGAGACCTTCATC	1925
525	QY	-----CysAlaValPheValValAsnGluThrAlaGlyAsnPro	537
1926	DB	TTGACTATCCCATGCATTTCTGGGCATCTTTTTTTTTTACCAAGAGAGCAATGGA-----	1979

Qy	538	CysLeuGluGluAenAspValValLysPheTyrThrGluLeuIleGlyCysLysPhe	557
Db	1979	-----	1979
Qy	558	ArgGlyIleArgIleGlyAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAla	577
Db	1980	CGAGCTCGAGAACTGGT-----CAACATGTTGGAGAAGATAGC	2018
Qy	578	ThrLysAenGluTyrAlaPheTyrLysAenCysThrLeuAenThrGlyIleGly-ArgPhe	597
Db	2019	CGGCCCCATGGTCATCGTATGAGCC-----ACCGCGCTGGTGTGAAC	2063
Qy	597	eGluIleAlaAlaThrGluAla---LysAenMetPheGluArgLeuProAspLysGluG1	616
Db	2064	AAAGATGACCGAATAGACATTTATGTCAGAACCATTCATCCACGCTTAGGAGCTGAGGG	2123
Qy	616	nLysValLeuMetPheIle---IleIleSerLysArgGlnLeuAenAlaTyrGlyPheVa	635
Db	2124	GAAGATACAGATGGTGTGTCATCATCATCGGCCCCACGTGATGATCTCTATGGGCCCAT	2183
Qy	635	lLysHisTyrCysAspHisThrIleGlyValAlaAenGlnHisIleThrSerGluThrVa	655
Db	2184	CAAGAACTGTGCTGTGTCAGTCCCAGTGCCTCCAGGTGTCAATGTTCCGAACCAT	2243
Qy	655	lThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnI1	675
Db	2244	TGTCACGCCACCAGG-----CTTCGAGATGTGGCCCCAGAACAT	2282
Qy	675	eAlaLeuLysIleAenAlaLysLeuGlyGlyIleAenGlnLeuAspTrpSerGluI1	695
Db	2283	TTTACTTCAGATTAACTGTAATTTGGTGT-----GAGCTC---TGGGGA-----	2325
Qy	695	eAlaGluIleSerProGluLulysGluArgArgLysThrMetProLeu-----ThrMe	713
Db	2326	-----GTGGATATTCCTCTGAAACAGATTAAT	2351
Qy	713	tTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAl	733
Db	2352	GGTGATCGGATGGATGTTTACCATGACCCCACT---AGAGGCATG---CGCTCCGTGGT	2405
Qy	733	aAlaValAlaSerIleAenProGlyThrIleTyrArgAenMetIleValThrG1	753
Db	2406	TGGCTTCGTGGCAAGCATCAATCTCACCTTCACAAAATGGTATPCCGGGTGGTGTCCA	2465
Qy	753	nGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLe	773
Db	2466	G-----ATGCCGCAT---CAGAGATTGTGGACAGCCT	2495
Qy	773	uGluAlaLysPheValLysLeuArgGluPheAlaGluAenAsnAspAenArgAlaPr	793
Db	2496	GAAGCTATGCTCGTGGCTCCTTAAAGAAAGTTTATGAGTGAAC---CACGTGCTACC	2552
Qy	793	oAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHi	813
Db	2553	AGAGAGATTGTGGTGACCGATGGTGTCTGATGCCAACTCAAGACAGCTGTCCAA	2612
Qy	813	sAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyG1	833
Db	2613	CTATGAGATTCTCAACTACAGAAGGTTTGAAGCTTT-----GA	2654
Qy	833	uAspGluProLysTyrThrPheIleValIleGlnLysArgHisAenThrArgLeuLe	853
Db	2655	GAATTATCACGCCCAAGATGGTGGTGTGTGTAGTTTCAGAAAGAAATCAGTACTAATCTA	2714
Qy	853	uArgArgMetGluLysAspLysProValValAenLysAspLeuThrProAlaGluThrAe	873
Db	2715	T-----	2715
Qy	873	pValAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrG1	893
Db	2716	-----CTGGCTGTCTCTCAGAACCTTT-----	2736

QY 893 yileValAenProSerSerGlyThrValValAspLysLeuIleValSerLysTyrLysPh 913
 Db 2737 ----GTAACCTCCACTCCTGGAACTGTGTAGATCATACAAACAGCTGTGAGTGGT 2792
 QY 913 eAspPheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrTh 933
 Db 2793 GGATTTCTATCTTCTTGCCCATCATGTACGGCAGGCTGTGGCATTCCTACGCATTATGT 2852
 QY 933 rValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLe 953
 Db 2853 CTGTGTTCTCAACACCGCAACCTGAGCCCTGATCATATGACAGAGCTGACTTCAAACT 2912
 QY 953 uAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAl 973
 Db 2913 GTGCCACATGTACTGAAATTCGCTGGCACCATCAGATTCACAGCTCTCTTGAAGTATGC 2972
 QY 973 aHis 974
 Db 2973 CCAC 2976

RESULT 10
 US-10-104-047-1469
 ; Sequence 1469, Application US/10104047
 ; Patent No. 6943241
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. 6943241el full length cDNA
 ; FILE REFERENCE: HI-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1469
 ; LENGTH: 2579
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-1469

Alignment Scores:
 Pred. No.: 1,39e-22 Length: 2579
 Score: 288.00 Matches: 144
 Percent Similarity: 36.7% Conservative: 87
 Beat Local Similarity: 22.9% Mismatches: 207
 Query Match: 5.4% Indels: 132
 DB: 3 Gaps: 27

US-10-645-746-3 (1-1020) x US-10-104-047-1469 (1-2579)

QY 417 GluValHisGlu---LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAsp 435
 Db 924 GAAGAACATAAATTTCAAGAACAGTTTCCAAAGAACTAATAGTTTGTCTTACC 983
 QY 436 LysPheLeuLysArgAlaThrArgLysProHis---AspTyrLysGluAsnThrLeuLys 454
 Db 984 AGATATAACAATAAGACATACAGAGTGGATGATATATGATCGGACCAAGATCCCAAGACC 1043
 QY 455 MetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheValGluArgPheGlyLeu 474
 Db 1044 ACCTTTAAGAAAGCCGACGGCTCT-----GAAGTCAGCTTCTTAGAATACTACAGGAAG 1097
 QY 475 CysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeuVal 494
 Db 1098 CAATACAAACCAAGAGATCACCGAC-----TTGAAGCAGCCTGTCTTGGTC 1142
 QY 495 AsnSerValAsnGluGln-----ile 501
 Db 1143 ACCCAGCCCAAGAGAGGGGGCCCTGGGGGACACTGCCAGGGCCTGCCATGTCTATT 1202
 QY 502 LysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsn----- 516
 Db 1203 CCTGAGCTCTGCTATTCTTACAGGTCTTAACGTATGATAAAATCGTAATGATTTTAACGTGATG 1262

QY 517 -----ValValProGluLysGluLeuCysCysAlaValPhe 528
 Db 1263 AAAGACTTAGCGTTTCATACAAGACTAACTCCAGAGCAAAAG----- 1304
 QY 529 ValValAsnGluThrAlaGlyAsn-----ProCysLeuGluGluAsnAspValValLys 546
 Db 1305 -----CAGCGTGAAGTGGGACGACTCATTTGATTACATTAATAAAACCGATAATGTTCAA 1358
 QY 546 ----- 546
 Db 1359 AGGAGCTTCGAGACTGGGGTTTGAGCTTTGATTCCAACTTACTGTCTTCTCAGGAAGA 1418
 QY 547 ---PheTyrThrGluLeuIle-----GlyGlyCysLysPhe-----ArgGlyIle 560
 Db 1419 ATTTTGCAAAACAGAAAGATTCCCAAGGTGGAATAAATTTGATTACATTCACCAATTT 1478
 QY 561 ArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsn 580
 Db 1479 GCAGATTGGTCCAAAGAAACAAGAGGTGCACCATTAATTTAGTGTTAAGCCACTAGATAAC 1538
 QY 581 GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluLeuAla 600
 Db 1539 TGGCTGTTGATCTATACGCGA-----AGAAATTATGAAGCAGCC 1577
 QY 601 AlaThrGluAlaLysAsnMetPheGluArgLeuPro----- 612
 Db 1578 AATTCATTGATACAAAATCTATTAAAGTTACACCAGCCATGGCATGCAATGAGAAAA 1637
 QY 613 -----AspLys-----GluGlnLys 617
 Db 1638 GCAATAATGATTGAAGTGGATGACAGAACTGAAGCTACTTAAGAGTCTTTACAGCAAAAG 1697
 QY 618 Val-----LeuMetPheIleIleIleSerLysArgGlnLeuAsnAlaTyr 632
 Db 1698 GTCAAGCAGACACCCAGATAGTTGTCTGTCTGTCTCAAGTAATCGGAAGCAATAATAC 1757
 QY 633 GlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSer 652
 Db 1758 GATGCTATTAAAAAATACCTGTGTACAGATTCCTCCCTACCCCAAGTCAGTGTGTGGGCC 1817
 QY 653 GluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePhe 672
 Db 1818 CGAACCTTTAGGCAAA-----CAGCAAACTGTCTATGGCCATTGCT 1856
 QY 673 TyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTyr 692
 Db 1857 ACAAGATTGGCCTACAGATGAACCTGCAAGATGGGAGGA-----GAGCTC---TGG 1904
 QY 693 SerGluIleAlaGluIleSerProGluGluLysGluArgArgLysThrMetPro----- 710
 Db 1905 -----AGGCTGGACATCCCTCTGGAAG 1925
 QY 711 LeuThrMetTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyr 730
 Db 1926 CTCGTGATGATCGTTGGCATCGATTGTACCATGACATGACAGCTGGG-----CGAGG 1979
 QY 731 SerIleAlaAlaValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIle 750
 Db 1980 TCATCGCAGGAGTTTGTGCCAGCATCAATGAAGGGATGACCCGCTGGTTCACCGCTGC 2039
 QY 751 ValThrGlnGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThr 770
 Db 2040 ATATTTTCAGGAT-----AGAGCAGCAGGCTGGTA 2069
 QY 771 AspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsn 790
 Db 2070 GATGGCTCAAAGTCTGCTGCAAGCGGCTCTGAGGGCTTGGATAGCTGCAATGAG--- 2126
 QY 791 ArgAlaProAlaHisIleValValTyrArgAspGlyValSerSerGluMetLeuArg 810
 Db 2127 TACATGCCAGCCGATCATCGTGTACCGCATGCGTAGGAGACGCGCCAGCTGAAAAACA 2186

```
QY 811 ValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSer----- 828
Db 2187 CTG-----GTGAACCTACCAAGTGCACAGTTTGGATTGTCTA 2225
QY 829 GluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHis 848
Db 2226 AATCCATTGGTAGAGGTACCAACCTAGACTAAGGTAAATGGTGGTGAAGAAAGAGTG 2285
QY 849 AsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLysAspLeuThr 868
Db 2286 AACACCAGATTTTTT----- 2300
QY 869 ProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLysGlu 888
Db 2300 ----- 2300
QY 889 SerLysGluThrGly-----IleValAsnProSerSerGlyThrThrValAspLysLeu 906
Db 2301 ---GCTCAGCTCGGAGGAGACTTCAGAACTCCACTTCCTGGACAGTTATTGATGAGAG 2357
QY 907 IleValSerLysTyrLysPheAspPheLeuAlaSerHisGlyValLeuGlyThr 926
Db 2358 GTTACCAGACCAAGATGGTATGACTTTTATCTGTGAGCCAGGCTGTGAGAGGTGTAGT 2417
QY 927 SerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluVal 946
Db 2418 GTTCTCCACACATTACATCATCTATGACACACAGCGCGCTGAAGCCAGACCCACATA 2477
QY 947 TyrLysMetThrTyrGlyLeuAlaPheLeu 956
Db 2478 CAGGCTTGACCTACAGCTGTGCCAATC 2507
```

RESULT 11

```
US-10-104-047-1676
; Sequence 1676, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1676
; LENGTH: 3138
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1676
```

```
Alignment Scores:
Pred. No.: 2,54e-21 Length: 3138
Score: 278.50 Matches: 189
Percent Similarity: 33.4% Conservative: 144
Best Local Similarity: 19.0% Mismatches: 329
Query Match: 5.2% Indels: 335
DB: 40 Gaps: 40
```

US-10-645-746-3 (1-1020) x US-10-104-047-1676 (1-3138)

```
QY 40 ValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyr 59
Db 467 GTGAAACTGGTTACAAACCTCTTTAACTTAGATTTTCCC---CAAGACTGGCAGCTATAC 523
QY 60 GluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhePro 79
Db 524 CAGTACCATGTGCATATATTTCCAGATTTAGCATCTAGAGAGGCTG----- 568
QY 80 LysLysThrGluIleProLysProAspAlaLysLeuPheTrpGlnHis-----Leu 97
Db 569 -----AGNATGGCTTACTTTATAGTCATAGTGAACCTT 601
```

```
QY 98 ArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp 117
Db 602 TCCAAACGAAGCAAAAGCATTCGACGGTGCATC----- 634
QY 118 ThrValTyrSerValCysArgLysLeuAsnThrValThrSerLysMetLeuValSerGluLys 137
Db 635 -----CTTTTCTGTCTCAAAAG 652
QY 138 ValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysIleLeuTyr 157
Db 653 CTAGAAGAAAGGTCCACGAGTTGTCAAGTGAATACTCAAGAGGTGAGACTATATAAGATG 712
QY 158 ThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
Db 713 ACTATCACCCCTG-----AAGAGGGAGCTGCCATCA 742
QY 178 LysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysVal 197
Db 743 AGTTCTCCCGTGTGCATCCAGGCTCTC-----AATATCATCTTCAGAAAGATC 790
QY 198 ArgTyrAlaProPheValAsnGluLysValGlnPheAlaLysAsnPheValTyr 217
Db 791 -----CTCAAAAAGTTGTCCATGTACCATAATTGGACGGAACCTC---TAT 832
QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
Db 833 AATCCTTCAGACCAATGGAAATTCGCCAGCAC----- 865
QY 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGlu 257
Db 866 -----AAATTATCCCTTTGGCTGGTTTGCATTTCTGTG---TCA 904
QY 258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
Db 905 TATTTTGAAGGAAGCTCCGTTTAGTGTGATGTGAGTTACAAAGTCTCTGTCTTCCACCAG 1024
QY 278 ProLysMetSerLeuLeuAspTyrLeuLeu-----LeuIleValAspProGln 293
Db 965 ACGGTTCTGGAAATTCATGACTGCTCTCTGTCAAAGAACTGGCTGTCTGTCTTCCACCAG 1024
QY 294 SerCysAsn----- 296
Db 1025 ACGTGTGAGAAGCAGCTAATAGGGCTCATTTGCTCTTACAAGATACAATAACAGAACCTAC 1084
QY 297 -----AspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys----- 311
Db 1085 TCCATTGATGACATTCAGTGTGCTCAGTGAACCCACACACACACCTTTCAGAACGGGATGGC 1144
QY 312 -----MetThrIleArgGlnAla 317
Db 1145 ACCGAGATCACCTATGTGGATTACTACAAGCAGCAGTATGATATTAATCTGATCGACCTG 1204
QY 318 AlaArgProArgIleArgGlnLeuLeuGluAsnLysLysCysAlaGluValTrp 337
Db 1205 AATCAGCCCCATGCTTGTAGTCTGTTAAAGAAAGAGAGA----- 1243
QY 338 AspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGlu 357
Db 1244 ---AATGACACACAGTCAAGCTCAGCTCGGCCACCTG---ATACCTGAGCTCTGC----- 1291
QY 358 AsnSerLeuValTyrLysValThrGlyLysSerAspArgGlyArgGlyArgHisLysTyr 377
Db 1292 -----TTTCTAACAGGGCTGACTGACCAGGCAACATCT----- 1324
QY 378 AspThrThrLeuPheLysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeu 397
Db 1325 GATTTCCAGCTGATGAAGCTGTGCGCTGAAAGACACAGT----- 1363
QY 398 ProLeuValLysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGlu 417
Db 1364 -----CTCAGT 1369
```


US-10-645-746-3 (1-1020) x US-09-248-796A-6819 (1-771)

```
QY 714 TyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAla 733
Db 16 TACTTGTGCTGTGATGTTACT-----TCATCACAACAGGAGAAATTTATCTGAATCA 69
QY 734 AlaValValAlaSerIle---AsnProGlyGlyThrIleTyrArgAsnMetIleValThr 752
Db 70 GTTCTATTGCTTCTATGTTGGTAGGAGATGGAATTTTCAATAAATTC-----120
QY 753 GlnGluGlyCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIle 772
Db 121 -----CCTGCATCAGTTCTGATTCCAACTGGTGACAAGAGTT---ATT 162
QY 773 LeuGluAlaLysPheValLysLeuLeuArg---GluPheAlaGluAsnAsnAspAsnArg 791
Db 163 GCTGATGTTAAAGTAGTATGTTTGGACAGATTAGAAAATTTCCATAAAAAAATTTGGGAAA 222
QY 792 AlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgVal 811
Db 223 TTACCTAGTAAAGTATTATTATTCGTGATGGAGTCTCTGAAGGACAATATACACCCATA 282
QY 812 SerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAsp 831
Db 283 TTAAGAAGCAATTAACAAAATTAAGCCGCTTTTAATGAATATGTAATTAAGAAAAC 342
QY 832 GlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArg 851
Db 343 ATCCCTAAATATCTCCAAACCATATCATTTATGTTGTTAAAGACATCATACAAGA 402
QY 852 LeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGlu 871
Db 403 TTTATT-----CCTATTCTCATGATAAT-----423
QY 872 ThrAspValAlaValAlaValLysGlnTrpGluAspMetLysGluSerLysGlu 891
Db 424 -----GCCGATCATCCAAAACCAAAAACAA 450
QY 892 ThrGlyIleVal-----AsnProSerSerGlyThrThrValAspLysLeuIleVal 908
Db 451 ATAGCTGTCACCAAGTAATGAAAATGTCATGCTGTGTACAAACCGTTGATCGAGAAATCACT 510
QY 909 SerLysTyrLysPheAspPheLeuAlaSerHisGlyValLeuGlyThrSerArg 928
Db 511 TCTCCAGCATTTTTCATTTTACGTTTCAATCTCAACAATCATTAACAGGTACTGGAATA 570
QY 929 ProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLys 948
Db 571 CCAGCTCATATTATGTTTACATGATGAAAATAATTATATCTTCAGATACTATACAAAATA 630
QY 949 MetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProVal 968
Db 631 ATCACTTATGATTATGTCATCTATTTTAGTAGAGTACTATAATCAGTCAAAAGTTGTCCCC 690
QY 969 ProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLys 988
Db 691 GCAGCTTATATGCCGATTATTATGTTACTAGAGGTAGAGAT-----732
QY 989 GluHisTyrIleGlyAspTyrAlaGln 997
Db 733 -----TATATTATGTTTGTCTAAA 753
```

RESULT 13

```
US-09-297-648-2528
; Sequence 2528, Application US/09297648
; Patent No. 6964868
```

GENERAL INFORMATION:

```
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
```

```
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassen, Altaf
; APPLICANT: Lanson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; TITLE OF INVENTION: Products II
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2528
; LENGTH: 734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(734)
; OTHER INFORMATION: n = A, T, C or G
US-09-297-648-2528
```

Alignment Scores:

Pred. No.:	7,61e-20	Length:	734
Score:	256.50	Matches:	57
Percent Similarity:	50.6%	Conservative:	23
Best Local Similarity:	36.1%	Mismatches:	47
Query Match:	4.8%	Indels:	31
DB:	4	Gaps:	3

US-10-645-746-3 (1-1020) x US-09-297-648-2528 (1-734)

```
QY 833 GluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeu 852
Db 118 AAAGACTATCAACCTGGAATACCTACATTTGTAGTTTCAGAGAGACATCACACTCGATTA 177
QY 853 LeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThr 872
Db 178 TTT---TGTCGTGATAGGACAGAAAGGGTTGGAAAGATGGCAATATCCAGCT-----228
QY 873 AspValAlaValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThr 892
Db 228 -----228
QY 893 GlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLys 912
Db 229 -----GGAACAACAGTTGATACAGACATTACACACCCATATGAG 267
QY 913 PheAspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyr 932
```

```

Db      268 TTGATTTTACCTCTGTAGCCATGCTGGNATACAGGGTACCAGTCTGCTTCACACTAT 327
Qy      933 ThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGly 952
Db      328 CATGTTTATGGGATGATAACTGCTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAG 387
Qy      953 LeuAlaPheLeuSerAlaArgCysArgLysProlIleSerLeuProValProValHisTyr 972
Db      388 CTCGTCACCACTTACGTCGCTGTACACGATCTGTTTCTATACCTGCACGCGTATTAT 447
Qy      973 AlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db      448 GCTCACTGTGACATTAGACCCACA---TATCATCTTGTGGACAAAGACAT 498

RESULT 14
US-09-873-737A-1
; Sequence 1, Application US/09873737A
; Patent No. 6723534
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE
; FILE REFERENCE: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/873, 737A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110,901
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3047
; TYPE: DNA
; ORGANISM: Drosophila sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(2612)
; NAME/KEY: misc feature
; LOCATION: (120)
; OTHER INFORMATION: n=a or c, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (399)
; OTHER INFORMATION: n=a or t, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (2436)
; OTHER INFORMATION: n=a or c, Xaa=Leu or Ile
US-09-873-737A-1

Alignment Scores:
Pred. No.:      9,96e-13      Length:      3047
Score:          205.00      Matches:      175
Percent Similarity: 33.7%      Conservative: 137
Best Local Similarity: 18.9%      Mismatches: 270
Query Match:      3.8%      Indels:      346
DB:               3         Gaps:      45

US-10-645-746-3 (1-1020) x US-09-873-737A-1 (1-3047)

Qy      231 AspProAsnArgPheGluGlnSerLeuGluValAlaProArgIleGluAla----- 247
Db      225 GAGGCTTCAAGAGAGAGAGAGAGCTCTCGAGGAAGCTCCCGAGCGGTGAGGTGGCGGCCA 284
Qy      248 -----TrpPheGlyIleTyr-----IleGly 254
Db      285 GAGCGAAAGCCGTGGGGTGCACCAATATGATTACCTGAATACCGTCCGCTTGAGCTGGTA 344
Qy      255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
Db      345 TCCAAGAAGGAAGCAACGATCGCGTCCGGTCATG-----CTGCAGACGAACATT 392
Qy      275 TyrAsnAla-----ProLysMetSerLeuAspTyrLeuLeuIleValAsp 291

```

```

Db      393 TTTTCANTAAACCAACGCGGAATGGCGGATCGTTTCATTATCAGCTGGAGTTGTGCGG 452
Qy      292 ProGlnSerCysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
Db      452 ----- 452
Qy      312 MetThrIleArgGlnAlaAlaArgProArgIleArgGln----- 324
Db      453 ---ACCATCGAG---AATCCTCGTGTCCGTATGGGAGTTTGTGCCAATCATGCT 500
Qy      325 -----LeuLeuGluAsnLeuLysLeuLysCysAlaGluValTyr 337
Db      501 AACCTTCTGGGATCAGGCTATCTATTCGACGCACTGCAACTGTTACACACAGGAATTC 560
Qy      338 AspAsnGluMetSer-----ArgLeuThr 345
Db      561 GAGCAGGAAATCACCGTGTCTCAGCGAAAGTCGAAAGTCGACATTTGAATACAAGATATCC 620
Qy      346 GluArgHisLeuThrPheLeuAspLysCysGluGluAsnSerLeu-----ValTyrLys 363
Db      621 ATAAAGTTTGGTGGATTTCATATCGTGTCTGAGCCCGCTTTTGCAGATCTTTAAATCTA 680
Qy      364 ValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLys 383
Db      681 ATATTGCGCGCTCGATGAGGCCCTAAATTGGAAATTAGTTGGCCGTAATCTCTTTGAT 740
Qy      384 IleTyrGluGluAsnLysPheIleGluPheProHisLeuProLeuValLysValLys 403
Db      741 -----CCCGAGCTAAGATCGAA 758
Qy      404 SerGlyAlaLysGluTyrAlaValProMet-----Glu 414
Db      759 -----ATAAGGAGTTCAAAATGAGCTATGCGCGGCTATGAGACATCGATTCGTCTAG 812
Qy      415 His-----LeuGluValHisGluLysProGlnArgTyrLysAsn 427
Db      813 CACGAAAGAATATTTTATTTGGCCACCGAAATAACTCACAAAGTTATGCGCACCGAGACG 872
Qy      428 ArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAsp 447
Db      873 ATCTACGACATA-----ATGCGACGTTGCTCACACAAATCGGTCGT 914
Qy      448 TyrLysGlu-----AsnThrLeuLysMetLeuLysGluLeuAspPheSerSer 463
Db      915 CATCAGGACGAAGTACGGGTAATGTTTGGACTTGATTTGCTTACGGATTACAATAAC 974
Qy      464 Glu-----GluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIle 481
Db      975 AGAACTTATCGTATCAATGATGTCGAC---TTTGGACAAACTCCGAAA---TCAACATTTC 1028
Qy      482 GluCysProGlyLysValLeu----- 488
Db      1029 AGTTGCAAGGGTAGAGATATAGTTTCGTGGAATACTATCTCACTAAATATAATATACGC 1088
Qy      489 -----LysGluProMetLeuValAsn----- 495
Db      1089 ATTCCGACCAACATCAGCGCGCTGCTGATTTCCAAAAATAGGACACAGGCTCTAAAAA 1148
Qy      495 ----- 495
Db      1149 AACGCTAGCGAATTAGTGTACTAATTCTCTGAGCTCTGCCGAGTGACTGGGCTCAATGCC 1208
Qy      496 -----SerValAsnGluGlnIleLysMetThrPro 505
Db      1209 GAGATGCGCTCAAACTTTTCAGCTTATGCGTGCCATGAGCAGTTATACGCAATGAACCCC 1268
Qy      506 Val-----IleArgGlyPheGlnGluLysGlnLeuAsnValValProGlu 520
Db      1269 AAACACGCACTGATCGATTGCGCGCTTTTAACCAACCGTTTACAAAAC---ACTCCAGAA 1325
Qy      521 LysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGlu 540

```

1326 -----AGTGTGAAGGCTCTTGAGAGAC-----TGAACATGGAACCTGGAC 1364
Db
541 GluAsnAspValValLysPheTyrThrGluLeuIleGly-----GlyCysLysPheArg 558
QY
1365 AAGAAC---GTCAACAGAGTCAACAGGCGCGGATATATGACAGCAGAACATCGTGTTCAT 1421
Db
559 GlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThr 578
QY
1422 AATGGAAGGTTCTGCTGGAGAAACCGTATTTGGCAAGGCACTTCAGAGCAAAAGG 1481
Db
579 LysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGlu 598
QY
1482 -----ATGCTTACCACCTCCGAGCGATGCGCTCGATCGTGGCT 1520
Db
599 IleAlaAlaThrGlu----- 603
QY
1521 GTCATCGCGCGCAAGAAATTCCTCAACTCCGAACCTCTACTTGACTCTTTGTATAGA 1580
Db
604 ---AlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysVal----- 618
QY
1581 GCAGCTAGTGAATGGGTCTTAGAATTCGAAGCCCGAGCAATTCATAATTTATGATGAT 1640
Db
619 -----LeuMet 620
QY
1641 CGCACTGGAACCTTATGTGAGAGCAATGATGATTTGTGTGCGCTCAGATCCCAACTTATA 1700
Db
621 PheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis-----Tyr 638
QY
1701 TTATGCTCTGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1760
Db
639 CysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVal---ThrLys 657
QY
1761 GTTGACAGG-----GCGGTGCCAATCAAGTTGTGACCCCTTAAACGACCAAGAACCGT 1814
Db
658 AlaLeuAlaSerLeuArgHisGluLysArgLysSerLysArgIlePheTyrGlnIleAlaLeu 677
QY
1815 AGCCTTATGAGC-----ATTGCCACCAAAATAGCAATC 1847
Db
678 LysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGlu 697
QY
1848 CAACTGAATTCGAATGGGA----- 1868
Db
698 IleSerProGluLysGluArgArgLysThrMetProLeuThr-----MetTyrVal 715
QY
1869 TATACACCTCGATGATCGAA-----CTACCCCTGTCCGAGCTGATGACAAT 1916
Db
716 GlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaVal 735
QY
1917 GGCTTTGACATTGCGAAGACACACGAGATCGGAGAGGCGCTAC-----GGAGCATTG 1970
Db
736 ValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGlu 755
QY
1971 ATTGCTCTCAATGGAT-----CTACAGCAAACTCCAC-GTACTTCAGCAC--- 2014
Db
756 CysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThr-AspIleLeuGlu-- 774
QY
2015 -----AGTCAGGAGTGCAGCGCCTTTGATGTGCTGCTGCTAA 2050
Db
775 -----AlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnAr 791
QY
2051 CACCCCTTTGCGCGATGATAGCAAGCCCTGCGCCCAATAT---CAACATGAGCATAGGAA 2107
Db
791 GalProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgVa 811
QY
2108 GCTGCCATCTCGAATCGTATTTTATCGAGACGGTGTGAGCTCGCGCTCTCTAAAGCAGCT 2167
Db
811 lSerHis-----AspGluLeuArgSerLeuLysSerGluValLysGlnPhe 827
QY
2168 TTTTGAATTTGAAGTCAAGGACATCATTTAGAGAGTTGAAGAACTGAATACGCCGCTCCA 2227
Db
827 tSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysAr 847
QY
2228 GCTAAGC-----CCACCGCAATTAGCTATATTTGTGTGTAACCAAGATC 2269
Db

QY 847 gHisAsnThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLe 867
Db 2270 CATGAACACGCGCTTCTTCTCTCAACGACAA----- 2300
QY 867 uThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLy 887
Db 2300 ----- 2300
QY 887 sGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuI 907
Db 2301 -----AATCTCTCCGCTGGTACTATAGTTGATGATGATGAT 2335
QY 907 eValSerLysTyrLysPheAspPheLeuAlaSerHisGlyValLeuGlyThrSe 927
Db 2336 AACTCTGCCGAGAGATACGACTTTTATCTGGTCTCGCAACAGTTCTGTCAGGTTACAGT 2395
QY 927 rArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTy 947
Db 2396 GTCGCCGACAGCTACCAATGTTCTTTATAGCAGCATGGGTTCTCACCGGAGAAAATGCA 2455
QY 947 rLysMetThrTyrGlyLeuAlaPheLeu-----SerAlaArgCysAspLysPr 963
Db 2456 AAACTTACGTACAGATGTCCTTGTACTACTACTACTACTACTACTACTACTACTACT 963
QY 963 olleSerLeuProVal 968
Db 2516 AGC-AGTTTCCAGTA 2530
RESULT 15
US-09-270-767-13282/c
; Sequence 13282, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13282
; LENGTH: 2167
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-13282
Alignment Scores:
Pred. No.: 7,02e-10 Length: 2167
Score: 178.50 Matches: 70
Percent Similarity: 39.0% Conservative: 43
Best Local Similarity: 24.1% Mismatches: 91
Query Match: 3.3% Indels: 87
DB: 3 Gaps: 11
US-10-645-746-3 (1-1020) x US-09-270-767-13282 (1-2167)
QY 709 MetProLeuThr-----MetTyrValGlyIleAspValThrHisProThrSerTyrSer 726
Db 2154 CTACCCCTTGTCCGAGCTGATGACAATTTGACATTTGCGAAGACACACGAGATCGG 2095
QY 727 GlyIleAspTyrSerIleAlaValAlaSerIleAsnProGlyGlyThrIleTyr 746
Db 2094 AAGAGGCGCTAC-----GGAGCATTTGATGCTCAATGAT-----CTACAG 2053
QY 747 ArgAsnMetIleValThrGlnGluCysArgProGlyGluArgAlaValAlaHisGly 766
Db 2052 CAAACTCCAC-GTACTTCAGCAC-----AGTCACGGA 2021
QY 767 ArgGluArgThr-AspIleLeuGlu-----AlaLysPheValLysLeuLeuAr 782
Db 2020 GTGCAGCGCTTTGATGTGCTCGCTAACCCCTTTGGCCGATGATAGCAAGGCCCTTGG 1961

```
QY 782 gGlupPheAlaGluAsnAsnArgAlaProAlaHisIleValValTyrArgAspG1 802
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1960 CCAATAT---CAACATGAGCATAGGAAGCTGCCATCTCGAATCGTATTTATCGAGACGG 1904
QY 802 yValSerAspSerGluMetLeuArgValSerHis-----AspGluLeuArgSe 818
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1903 CGTGAGCTCGGCTCTCTAAGCAGCTTTTGAATTTGAAGTCAAGGACATCATTCAGAA 1844
QY 818 rLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLy 838
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1843 GTTGAATACTGAATACGCCCGCTCCAGCTAAGC-----CCACCGCA 1802
QY 838 sTyrThrPheIleValIleGlnLysArgHisAenThrArgLeuLeuArgArgMetGluLy 858
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1801 ATTAGCTTATATTGTTGTAACAGATCCATGAACACGCGCTTCTCCTCAACGGACAA-- 1744
QY 858 sAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAlaAl 878
Db ----- 1744
QY 878 aValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleValAsnProSe 898
Db -----AATCCTCC 1736
QY 898 rSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPheLeuAl 918
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1735 GCCTGTACTATAGTTGATGACGTTATAACTCTGCCCGAGAGATAGACTTTTATCTGTT 1676
QY 918 aSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAs 938
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1675 CTCGCAACAAGTTCGTCAGGGTACAGTGTCCGCGACCAGCTACAATGTTCTTTATAGCAG 1616
QY 938 pLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeu---- 956
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1615 CATGGTCTCTCACCGGAGAAAATGCAAAACTTACGTACAAAGATGTGCCACTTGTACTA 1556
QY 957 -----SerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHi 974
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1555 CAATTGGTGGGACCACACGAGTGCCA----- 1528
QY 974 sLeuSerCysGluLysAlaLysGluLeu 983
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1527 -GCCGTTTCCAGTACGCTAAGAAGTTA 1501
```

Search completed: July 5, 2006, 23:28:20
Job time : 471 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:46:42 ; Search time 23 Seconds
(without alignments)
1189.943 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 MMSNPBLKGFVRLSDPE.....RHMEHPLQTNVYKPGMSFA 1020

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112942 seqs, 26832045 residues

Total number of hits satisfying chosen parameters: 112942

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pap.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	577	10.8	857	US-11-192-437-1	Sequence 1, Appli
2	566	10.6	859	US-11-192-437-2	Sequence 2, Appli
3	555.5	10.4	860	US-11-192-437-3	Sequence 3, Appli
4	555	10.4	861	US-11-192-437-4	Sequence 4, Appli
5	535	10.0	1050	US-10-953-349-4064	Sequence 4064, Ap
6	531	9.9	870	US-10-953-349-4065	Sequence 4065, Ap
7	509	9.5	1052	US-10-449-902-53188	Sequence 53188, A
8	507.5	9.5	822	US-10-953-349-4086	Sequence 4086, Ap
9	500	9.3	1011	US-10-449-902-56159	Sequence 56159, A
10	431.5	8.1	876	US-10-449-902-37356	Sequence 37356, A
11	427	8.0	1145	US-10-997-086-5	Sequence 5, Appli
12	421.5	7.9	517	US-10-449-902-38630	Sequence 38630, A
13	403.5	7.5	457	US-10-953-349-18475	Sequence 18475, A
14	403.5	7.5	458	US-10-953-349-18474	Sequence 18474, A
15	403.5	7.5	501	US-10-953-349-18473	Sequence 18473, A
16	396.5	7.4	527	US-10-953-349-19398	Sequence 19398, A
17	370.5	6.9	277	US-11-192-437-40	Sequence 40, Appli
18	370.5	6.9	488	US-10-953-349-19399	Sequence 19399, A
19	366	6.8	630	US-10-449-902-47718	Sequence 47718, A
20	361.5	6.8	471	US-10-953-349-19400	Sequence 19400, A
21	352.5	6.6	277	US-11-192-437-41	Sequence 41, Appli
22	344.5	6.4	277	US-11-192-437-39	Sequence 39, Appli
23	341.5	6.4	287	US-11-192-437-42	Sequence 42, Appli
24	302	5.6	292	US-10-449-902-54734	Sequence 54734, A
25	284	5.3	266	US-10-449-902-31632	Sequence 31632, A

26	284	5.3	256	6	US-10-449-902-32271	Sequence 32271, A
27	260.5	4.9	666	7	US-11-293-697-3968	Sequence 3968, Ap
28	234.5	4.4	763	6	US-10-449-902-43163	Sequence 43163, A
29	156	2.9	3113	6	US-10-505-928-325	Sequence 325, App
30	151	2.8	1328	6	US-10-504-973-32	Sequence 32, Appl
31	143.5	2.7	1502	6	US-10-471-571A-1696	Sequence 1696, Ap
32	142	2.7	956	7	US-11-293-697-3037	Sequence 3037, Ap
33	140	2.6	2871	6	US-10-505-928-100	Sequence 100, App
34	139	2.6	1354	7	US-11-270-653-1	Sequence 1, Appli
35	138	2.6	1674	6	US-10-511-937-2587	Sequence 2587, Ap
36	136	2.5	108	6	US-10-449-902-32719	Sequence 32719, A
37	136	2.5	436	6	US-10-471-571A-5206	Sequence 5206, Ap
38	133.5	2.5	1116	6	US-10-449-902-47106	Sequence 47106, A
39	129	2.4	824	7	US-11-293-697-4573	Sequence 4573, Ap
40	127.5	2.4	1003	6	US-10-471-571A-4826	Sequence 4826, Ap
41	126.5	2.4	1517	7	US-11-257-500-9	Sequence 9, Appli
42	126.5	2.4	1526	7	US-11-257-500-15	Sequence 15, Appli
43	125.5	2.3	972	6	US-10-480-962-4	Sequence 4, Appli
44	124.5	2.3	94	7	US-11-192-437-13	Sequence 13, Appli
45	124.5	2.3	94	7	US-11-192-437-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-11-192-437-1
; Sequence 1, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-192-437-1

Query Match	10.8%	Score 577, DB 7, Length 857;
Best Local Similarity	23.0%;	Pred. No. 3e-28;
Matches 227;	Conservative 171;	Mismatches 365; Indels 222; Gaps 37;
QY	38	KKVLLVNNFKSSKIYDREYVEYKMTKEVLNRPKPGPKKTEIPIDRAKLFQHL 97
Db	34	KPKLLANFYFVDIPKIDVHYEVDIKPKD-----CPRVREVVE-----YMW 77
QY	98	RHEKQKTDPILEDY--VFDEKDTVSVCLNTVTSOMLVSEKV---VKQSEKKDKDLE 152
Db	78	QHFKPQ---IFGDRKPVYDGKNIY-----TVALPIGNERVDFVETIPGECKD-RIFK 127
QY	153	KKILYTMITYTKKFLNFSRENPEKDEANRSYKFLKNVMTOKRYAPVFNVEIKVQFA 212
Db	128	VS1KKWLAIYSWRMLHEALVSGQIPVPLESVQALDVAHRHLAS--MRYTPV----- 175
QY	213	KNFVYDNNLSILRVPSFHDNPFQESLEVAPEATEAMFGIYIGIKELFDGEPVLNPAVDK 272
Db	176	-----GRSFFSPPEGYHP-----LGGGREGVWFGHQSVRPAM-WKQMLNIDVSAT 220

QY 877 AAVQWBEEDMKESKETGIVNPSGTTVDKLVSKYKDFDFLASHGVLGTSRPHGYTMY 936
Db 733 -----GTTVDTKITHTEFDYFLCSHAG:QGTSPRSHYHLW 769
QY 937 DDKGMSQDEVYKMTYGLAFLASRCRKPISLPVPVHYAHLSCAKELVRYTKEH 990
Db 770 DNRFSDELQILTYQLCHTVTRSVSIPAPAYAHVAFRAR-YHLVDKEH 822

RESULT 3
US-11-192-437-3
; Sequence 3, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: ARGONAUTE PROTEINS
; CURRENT APPLICATION NUMBER: US/11/192,437
; PRIOR FILING DATE: 2005-07-28
; PRIOR FILING DATE: 2004-07-28
; PRIOR FILING DATE: 2004-07-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-192-437-3

Query Match 10.4%; Score 555.5; DB 7; Length 860;
Best Local Similarity 23.2%; Pred. No. 6.5e-27;
Matches 230; Conservative 149; Mismatches 377; Indels 237; Gaps 35;

QY 38 KKVLVNVNFKFSKLYDRYYEYVQWTKVNLNRPKPPKTEIPIDRAKLFQWHL 97
Db 28 KPIKLLANCFQVBIPIKIDVLYEVDIKPKD-----CPRRVNREVVD5---MVQHF 74
QY 98 RHEKQOTDILEDYVDFEVDKTVYVSVCLNTVTSMVLSEKVKVSKDSEKKDKLEKILY 157
Db 75 ----KVTIFGDRPVVDGKSLTANPLPVATTG-----VLDVTLPGEKGKDRPFKV 123
QY 158 TMLTYRKXPHLNFSENPEKBEANRSYKFLKNVMTQKRYAPFVNEEIKVQFAKNFY 217
Db 124 SIKFVSRVSWHL-----LHEVLGRTLPEPL---ELDKPISNPNVH 161
QY 218 DNNSILR-----VPSFHPDNPFBQSLEVAPEIAWFGIYIGIKBLFDG 261
Db 162 AVDLVRLHPSMKYTPVGRSFFSAPGEGYDHP-----LGGREVMFGFHQSVPAM-W 212
QY 262 EPLVNEAIVDKLFYNAPKNSLLDYLLIVD-----PQSCNDVDRKDLTKLWACKMTI 314
Db 213 KMLNLDVSNATFYKA--QPVIOFMCEVLDIHNDIQPRPLTDHVRKFTKEIKGLKVEV 270
QY 315 QMARPRIRQLLENLKLKCAEVWDNEMSLRTERHLTF-LDLCEBSLVYKVTGKSDRGRN 373
Db 271 THCGTWRRKRYVCN-----VTRRPAHQTFPLQL-----EN 301
QY 374 AKKYDTTLFKIYEENKFP-IEFPHLPLVKKVSAKAYVPMHELVHEKRPQRYKNRIDLV 432
Db 302 GQTVERTVAYQFREKTYLQKYPHLPCLQVGQKQKHTYLPLEVNT-VAGQRCIKLTDN 360
QY 433 MQDKFLKRAIRKPHDHYKENTKMLKELDFSESSELNFEVERFGLCSKLQWIECPKVLKEM 492

Db 361 QTSMIKATARSAPDROEBISRLVRSANYETDP--FVQEFQFKVRDEMAHVTGRVLPAPM 418
QY 493 LVNSVNEQIKMTVP-----IRGFOEKQLNV-VPEKELCCAVFVNVNETAGNCLBENDVVK 546
Db 419 LQYGRNRRTVATPSHGVDMRG---KQFHTGVEIKWALACFATQ-----RQCRBE--ILK 469
QY 547 FYTELI-----GCCKFRGIRIGANENRGAQSIYDATIONEYAFYKXNCTLNTGIGRPEIA 600
Db 470 GFTDQLRKISKDAGMPIQGOFCCKYAQGADSV-----502
QY 601 ATEAKMFERLPDKEQKVLMIIFIIISKQLNAYGFVHKYCDHTIGVANOHLITSETVTYKALA 660
Db 503 -----EPWFRHLKNTYSGLQILVILPQKTPVYAEVRKVGDTLLGNATOCVQKVKVKT--556
QY 661 SLRHEKSGKRIFYQIALKINAKLGGINOELDWSEIAEISPEEKERRKRTMPLTMVYGVDDVT 720
Db 557 -----SPQTLNLCIKINVLGGINNIL-----VPHQPSVFPQPV-IFLGADVT 600
QY 721 HPTSYSGIDYIAAVVASINPGGTIYRMTIVTQECRPPGERAVAHGRERTDILE--AKFV 778
Db 601 HPPAGDGKPKSIAAVVGSMDAHPFSRYCATVRVQ-----RPROBIIQDLASWV 647
QY 779 K-LLRSPAENNDNRAPAHIVVYRDGVSDEMLRVSHDELRLSKSEVKQPMSERDGEDPEP 837
Db 648 RELIQQPKSTRFK-PTRIIFYRDGVSEGFQVLYELLALREACISL-----EKYQ 701
QY 838 KYTFVIQKRNTRLRRMEKDKPVVNVKOLTPAETDVAVAVKQWBEEDMKESKETGIVNP 897
Db 702 GITVIVVQKHHHTLP-CADRTERRVGRSGNIPA-----733
QY 898 SSGTTVDKLVSKYKDFDFLASHGVLGTSRPHGYTMYDDKMGMSQDEVYKMTYGLAPLS 957
Db 734 --GTTVTDITHPYEFDFYLCSHAG:QGTSPRSHYHLWDDNCFTADELQILTYQLCHTY 791
QY 958 ARCKKPISLPVPVHYAHLSCAKELVRYTKEH 990
Db 792 VRCRVSISIPAPAYAHVAFRAR-YHLVDKEH 823

RESULT 4
US-11-192-437-4
; Sequence 4, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: ARGONAUTE PROTEINS
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR FILING DATE: 2004-07-28
; PRIOR FILING DATE: 2004-07-28
; PRIOR FILING DATE: 2004-07-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-192-437-4

Query Match 10.4%; Score 555; DB 7; Length 861;
Best Local Similarity 23.0%; Pred. No. 7e-27;
Matches 227; Conservative 161; Mismatches 378; Indels 220; Gaps 36;

QY 38 KKVLVNVNFKFSKLYDRYYEYVQWTKVNLNRPKPPKTEIPIDRAKLFQWHL 97

Db 688 VDSKICHPTEFDYLCSHAGIQTSPAHYHVLWENNFTADGLQSLTNLCYTYARCT 747
QY 962 KPISLPVVPVYAHLSCEKAKELYRTYKEHYIGDYAQPT 1000
Db 748 RSVSIVPPAYYAAFRAR-----FVMEPET 774

RESULT 9
US-10-449-902-56159
; Sequence 56159, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 56159
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-56159

Query Match 9.3%; Score 500; DB 6; Length 1011;
Best Local Similarity 22.0%; Pred. No. 2.3e-23;
Matches 224; Conservative 172; Mismatches 384; Indels 236; Gaps 41;

QY 49 FSSKIYDREYVEYVMTKEVLRKPKGPPKTEIPDRAKLFWHLRHEKKQTFIL 108
Db 163 FFAHLPNKLHLYDVDSITPEVTSR-----INRAVI-----KELVNLX 201
QY 109 EDVY-----FDEKDTVYSCRLNTVTSKMLVSEKVKVKKSEKKKLEKKILYTWILT 162
Db 202 ASVLGRLPAYDGRKSLYTAGLPDFTSQEQFQIT---LLDDDDGGSGRRQRTFRVVIKFA 258
QY 163 YRKXPH---INFSENPEKDEANRSYKFLKNVMTQ--KVRYPAPFVNEEIKVQFAKNFVY 217
Db 259 ARADLRHLFLAGRAHAEPQEA--LQVLDIVLRELPSARYAPFGRSPFSAYLGR--- 311
QY 218 DNNILRVPEFHDPNRFQOSLEVAPRIEAWFGIYIGIKELFDGEPVLPFAIVDKLFYNA 277
Db 312 -----RQP-----LGEGLSEWRGFGYQSIPTQMGSL-LNIDMSATAFIE- 349
QY 278 PKMSLDYLLLIIVDPQSCNDVVRKDKLTKLMACKMTIROAARPIRQLLENLKLCAEVW 337
Db 350 -PLPVLDIV-----AQLNSDIH-----SRPLSDAERVKIKKALRGVKEVTH-- 391
QY 338 DNMSRLTERHLTFLDLCEANSIYKYVTGSKDRNAKKYDTTLFKIYBENKXF-IEFPH 396
Db 392 RGNMRR--KYRISGLTIQPTRELTFFV-----DEGGTVK-----SVVQYFQETYGAIQHTY 441
QY 397 LPLVKVSKAKAYAVPMHEHLEVHEKPORYKNRIDLVMDQKFLKRAKRPKHDKYKENTLKM 456
Db 442 LPCLTVQ---RLNLYPMVECKIVE-GQRYSKRLNQIRALLBETQCHPRDRERDIKKV 497
QY 457 KELDFSSBELFVERGLCSKLMQIECPGKVLKEPML-VNSVNEQIKMTPVI---RGFQE 512
Db 498 KNNAYQDDP--YAKEFGIKISDLASVEARILPAPRLKYNETREKDCPLRVQGMNMNMK 555
QY 513 KQLVNVPPEKELCCAVFVWNETAGNCPLENDVVVKFVTELIGCGKPRGI-----RIG 563
Db 556 KWVNGKVRVSWMVCNFAFN-----VQESVVRGFCHEALMCAQSGMDFAPFPIPLPLN 608

QY 564 ANENRGAQSTMYDATKNEYAFYKNCINTLTGIGRPEIAATEAKNPFELPDKEQKVLFI 623
Db 609 AHPDQ-----VERALKARY-----HDMNVLG--PORRELDLIGI 642
QY 624 ISKRQLNAVGFVKHYCDHTIGVANQHTSETVTKALASLRHKGSKRIFQIALKINAKL 683
Db 643 LPDNGSLYGDLLKRVCEIDLGIIVSCCCTKQVFM-----NKQILANLALKINVKV 693
QY 684 GGINQELDWSEIAEISPEEKERRKTWPL-----TMVVGIDVTHPTSYSGIDYSIAAVAS 738
Db 694 GGRNTVL---VDAYS-----RRIPLVTRPTIIFGADVTHPHPCGEDSSPSIAAVAS 742
QY 739 IN-PGCTIYRNMIVTQECRCPGERAVHGRE-----RTDILAEKFKVLREP 784
Db 743 QDMPEVTKYAGLSAQ-----AHRQELIEDLYKIWDPOQRTVSGGMIRELLISF 792
QY 785 AENNDRAPAHIVVYRDGVSDEMLRVSHDELSLSEKSVQFMQFMSERDGEPEPKYTFIVI 844
Db 793 KRSTGEK-PQRIIFIRYRDGVSQFYQVLLYELNARKACASLET-----NYQKATFIV 846
QY 845 QKQHNTLLRRMEKDKPVVKNKDLTPAETDVAAVAKQWEEDMKESKETGINVPSSGTTVD 904
Db 847 QKQHNTLLFAHNNDQNSVDR-----SGNILP--GTVVD 878
QY 905 KLIVSKYKDFFLASHHGVLTGSRPGHYTWYDDKMSQDEVYMTVGLAFLSARCKPI 964
Db 879 SKIRHPTFEDFYLCSHAGIKGTSRPAHYHVLWENNFTADALQILTNNLCYTYARCTRSV 938
QY 965 SLFPVPVYAHLSCEKAKELYRTYKEHYIGDYAQ-----PRTRHEMEHFLQTNVKYPG 1016
Db 939 SIVPPAYYAAHAAFR-----RFYNEPDTSDSSVSGVGVGRGLSGSSTRTAPG 990

RESULT 10
US-10-449-902-37356
; Sequence 37356, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 37356
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-37356

Query Match 8.1%; Score 431.5; DB 6; Length 876;
Best Local Similarity 22.1%; Pred. No. 3.5e-19;
Matches 223; Conservative 140; Mismatches 404; Indels 241; Gaps 37;

QY 26 RPTGKCDGKDYKFKVKKVLLLVNWFKFS-SKIYDREYVEYVMTKEV-LNRKPGKPPFKKTE 83
Db 35 RPNLSEGMIGE-SCIVRTNCFSVHLESDDQTIYEYDVCTVEGINRAVIRELVK--- 90
QY 84 IPIPRAKLFWHLRHEKKQTFILEDYVDEKDTVYSCRLNTVTSKMLVSEKVKKOS 143
Db 91 -----QKXDSGLGRLPAYDGRKRLYTSGLPFDHRLFLVLLDSIEDSP 134
QY 144 EKKDEKDEKLLKLYTMILTYRKFKHL-----NESRENPEKDEANRSYKFLKNVMTQ--KV 197
Db 135 EBSRLRVRD---FVTLKFAAKISLWLRKFRGGKPNRESRA--ALRALDVVLKELPTA 189

```

QY 198 RVAPFVNEEIKYQAKNFVYDNNLSILRVDPESPHDNRPSQSLVAPRIEAWFGVIGIKE 257
Db 190 RTQTAG-----SYSPN-LGECQOLCKVLESWEGFQRIQA 225
QY 258 LFDGPEVLNFAIVDKLFYNAPKMSLLDYLLIVDPQSCNDDVVRKOLTKMLAKGMTIRQA 317
Db 226 TQMGQLQ-LNIDVSSVFIK--PVPVVDV-----AQLLNEDILLD----- 262
QY 318 ARP-----RROLLENILKCAEVDWDMNSRLTERHLTFLDLCEENSLVYKVTGKSDR 370
Db 263 -RPLCSTEFKLKEALEGLKVQI-----NGILENTYHVQDL 297
QY 371 GRNAKYDTTLFKIYEENKKEFEFHLPLVKVKGAKVAVPMHELVHEKQVRKNR 430
Db 298 VHQASFPVNF-----SIQYSLPCLKVNAHGETIFLFEVCKKAE-CQHQKQNLN 347
QY 431 LVMDQKFLKRAKRPKPHYKENTLKMELDFSEELNFERFGLCLSKQMIKPCQKVLKE 490
Db 348 AKHMAALLQVAPQPNERNYDNIQTVHONKY--QEDPHAKEFGIKIEEKLVSIKSRIIPA 405
QY 491 PML-----VNSVNEQIKMTFVIRGFOEKQLNVVPEKELCCAVFV--VNETA-----GN 536
Db 406 PWLKFHDSGETTEFLPQLGIMNMHKOMINGGRVKSWACVNFCSVREYAARNFCYDLGF 465
QY 537 PCLLENDV--VKFYTELI--GCKFRGIRIGANENRGAQSIWYDATKNEYAFYKNCNLN 591
Db 466 MCRESGMFSVAPVPLVTAKEGCVESALR-----TLH 498
QY 592 TGIGRFEIATBAKMFERLPDKEQKVLMIIFIIISKRLNAYGFVKHYCDHTIGVANQHIT 651
Db 499 DDV-----MDILRPQGRKLDLILVILPNNNGSLYGDVKRICETDGLISQCC 546
QY 652 SETVTYKALASLRHEKSGRIFFQIQAALKINAKLGGINQELDWSEIAEISPEEKERRKTWPL 711
Db 547 AKHVLKM-----NWKYLASVALKINAKMGGRNTVL-----VDALERMRLPHVRDTP-- 591
QY 712 TMYVGIDVTHPTSYSGIDYSIAAVVASIN--PGTITYRNMIVTQ---EECRPGERAVAGR 767
Db 592 TIVGAHVTHPHGPGKANSSIAAVASQDWPEVTKYAGISVQAQCHQESIQGLFKVQDDP 651
QY 768 ERTDILEAKFVKLLREFANNDRNAPAHIVVYRDGVSDSEMLRSHDELRLSLKSEYKQPM 827
Db 652 ERGTTSGMIKEHLASFYBAT--KRKEGRIFFYRDGVSKQLPQALMHELGAIKMACASM- 709
QY 828 SERDEGDPKPTFVIQKRNHTRLLRMEKQKPVVKNKDLTFAETDVAVAVKQWDEDMK 887
Db 710 -----GPDYNPLVTVYVQLCRHTRLF-----ADYNNANTHDSANIR----- 747
QY 888 ESKETGINVPSSTVVDKLIIVSKYKDFDFLASHHGVLGTSRPHGTYVMYDDKGMQDEVY 947
Db 748 -----AGTVVDNSNICQNPQDFLCSHRSTQGTKRPRYTHVLMDENDFLAGSFQ 796
QY 948 KMTYGLAFLSARCRKPISLPVPVHVHAHLSCEKAKELYRTY--KEHYIGD 994
Db 797 ELTNVLCYTSATCTQSISVAVPVHVARLLSSRA-----RCYIKPRSIGD 840

```

```

RESULT 11
US-10-997-086-5
; Sequence 5, Application US/10997086
; Publication No. US20060135456A1
; GENERAL INFORMATION:
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Paddison, Patrick J.
; APPLICANT: Siolas, Despina C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: CSHL-P05-010
; CURRENT APPLICATION NUMBER: US/10/997,086
; CURRENT FILING DATE: 2004-11-23
; PRIOR APPLICATION NUMBER: US 10/350,798
; PRIOR FILING DATE: 2003-01-24

```

```

; PRIOR APPLICATION NUMBER: US 10/055,797
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/08435
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 09/866,557
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/189,739
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/243,097
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-997-086-5

```

```

Query Match      8.0%; Score 427; DB 6; Length 1145;
Best Local Similarity 21.1%; Pred. No. 9.8e-19;
Matches 210; Conservative 152; Mismatches 381; Indels 250; Gaps 37;

```

```

QY 59 YEYEVKMTYKVLNRKPGKPPKTIPIPDRAKLFQWHLRHEKKQTDPILEDYVDFEKT 118
Db 364 YHYDVKI-----MPEREKKEFYQAFEQFRVLDQLGGAVLAYDGKAS 403
QY 119 VYSVCR--LNTVTSKMLYSE---KVVKKDSEKKDEKD--LEKKILYTMILTYRKKPHLNF 171
Db 404 CYSVDKLPNLSQNPVTVTDNRGRTLRYTIEIKETGDSIDLUKSLTTYN----- 452
QY 172 SRENPEKDEEANRSYKFLKNVMTQKRVYAPFVNBEIKVQAKNFVYDNNLSILRVPESEF-- 229
Db 453 ---NDRIFDKPMRAMQCVENVLASPC-----HNKAIRVGRSPFK 488
QY 230 -HDPNRFQCSLEVAPRIEAWFIYIGIKELPDGPEVLNFAIVDKLFYNAPKMSLLDYLL 288
Db 489 MSDPNNRH---ELDDGYEALVGLYQAF--MLGDRPFLNVDI SHKSF--PISMPMIEYL-- 539
QY 289 IVDPOSCHDDVYRKDLTKLMAGKMTIRQAARPIRQLLENLKLKCAEYVDNEMSLRTERH 348
Db 540 -----ERFSLUKAKI--NNTNLDSYRRFLEPFLRGINV-----VTPPQS 577
QY 349 LTFDLCLBENSIVYKVTGSKDRGNAKKYD-----TTLFKIYEENKKEFEFHLPLVKYK 403
Db 578 -----FQSAPRVYRVNGLSRAPASSETFEHDKGKVTIASYPHSRNYPLKFPQLHCLNVG 631
QY 404 SKAGEYAVPMHELVHEKRPQRYKNRIDLVMQ--DKFLKRAKRPKPHYKENTLKMELKELDFS 462
Db 632 SSIKSILLPIELCSIEE--GQALNRKDGATQVANNMIKYAATSTNVRKRIKMNLLQYFOHN 689
QY 463 SEELNFVRRFGLCSKLQMECPGKVLKBPMLVNSVNEQIKMTPVIRGFQE---KOLNVV 518
Db 690 LDPT--ISRFGIRIANDFIVSTRVLSPPQV---BYHSKRFTWVXNGSWRMDGMKLEBK 744
QY 519 PEKELCCAVFV-----VNETA---GNPCLLENDVVKFYTELIGGCKFRGIRIGANEN 567
Db 745 PKAHKCAVLVCDPRSGRKNMYTQLNDFGNLIISQKAVNI--SLDSDVTVYRPF---TDDE 799
QY 568 RGAQSIMYDATKNEYAFYKNCNTLNTGIRGFEIATAEAKMFERLPDKEQKVLMIIFIIKSR 627
Db 800 RSLDTIFADLKRSQHD-----LAIVIPQF 824
QY 628 QLNAYGFVKHYCDHTIGVANQHITSETVTYKALASLRHEKSGRIFFQIQAALKINAKLGGIN 687
Db 825 RI-SYDTTIKQKAELOHQHILTOCIKQFTVER-----KCNNQTTGNILLKLNKLNGLIN 875
QY 688 OELDWSEIAEISPEEKERRKTMPLTMYVGIDVTHPTSYSGIDYSIAAVVASINPGGTIYR 747
Db 876 HKI-----KDDPRLPMKNTMYIGADVTHSPDQRETPSVVGVAAASHDPYCASYN 925
QY 748 NMIVQTEBCRPGERAVAGRERTDILEAKF---VKLLREFAENNDRNAPAHIVVYRDGV 804
Db 926 MQYRLQ-----RGALEIEDMFSTLEHLRVYKEYR-NAYPDHIIYYRDGVS 971

```



```
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18474

Query Match
Best Local Similarity 28.1%; Score 403.5; DB 6; Length 458;
Matches 121; Conservative 59; Mismatches 129; Indels 121; Gaps 14;

QY 612 PKQKRV-----LMFTIIISKRQLNAYGVKHYCDHTIGVANQHIITSE 653
Db 60 PDQVEKVLKTRYHDAKNKLQGRELDLLIVLPDNNGLSLYGLDKRICETDLGLVSOCLTK 119
QY 654 TTTKALASLRHEKSGKRIFYQTALKINAKLGGINQELDMSEIAEISPEEKERTMPL-- 711
Db 120 HVFKM-----SKQYLANVALKINVKVGGRTVL-----VDALS-----RRIPLVS 159
QY 712 ---TWVVGIDVTHPTSYSGIDYSIAAVASIN-PGCTIYRNMIVTQECPGERAVAHGR 767
Db 160 DRPTIIFGADVTHPHPGEDSSPSIAAVASQDYPEITKYAGLVCAQ-----VHRQ 209
QY 768 E-----RTDILEAKFVKLLREFAEENNDRAPAHIVVYRDGVSDSEMLRVSHD 814
Db 210 ELIQDLFKQWQDPVVRGTVTGGMIKELLISFRATGOK-PQRIIFYRDGVSEGQFYQVLLF 268
QY 815 ELRSLSKSEVKQPMSEKEDGDEPDKY-----TFIVIOKRHNTRLRLRMEKDKPVVVKDLTPA 870
Db 269 ELDAIRKACASL-----EPNYQPPVTFVVQKRHHTRLFASNHHDKSSVDK----- 314
QY 871 ETDVAVAVKQWEEDMKESKETGIVNPSGTTVDKLIYSKYKFPDFFLASHHGVLTGTSRPG 930
Db 315 -----SGNILP--GTVDYSKI CHPTFEFDLYLCSHAGIQGTSRPA 351
QY 931 HYTMVYDDKGMQDEVYKMTYGLAFLSARCKRPISLPVPVHYAHLSCSEKAKELYRTRYKEH 990
Db 352 HYHVLWDENNFTADALQTLTNLNCVYARCTRSVSIVPPAYYAHLAAPFRAR----- 402
QY 991 YIGDYAQPET 1000
Db 403 ---FYMEPET 409
```

```
RESULT 15
US-10-953-349-18473
; Sequence 18473, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18473
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18473

Query Match
Best Local Similarity 28.1%; Score 403.5; DB 6; Length 501;
Matches 121; Conservative 59; Mismatches 129; Indels 121; Gaps 14;

QY 612 PKQKRV-----LMFTIIISKRQLNAYGVKHYCDHTIGVANQHIITSE 653
Db 103 PDQVEKVLKTRYHDAKNKLQGRELDLLIVLPDNNGLSLYGLDKRICETDLGLVSOCLTK 162
QY 654 TTTKALASLRHEKSGKRIFYQTALKINAKLGGINQELDMSEIAEISPEEKERTMPL-- 711
Db 163 HVFKM-----SKQYLANVALKINVKVGGRTVL-----VDALS-----RRIPLVS 202
```

Search completed: July 5, 2006, 13:50:09
Job time : 29 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 14:11:38 ; Search time 23 Seconds
(without alignments)
1189.943 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 1020

Sequence: 1 MSSNPFLEKGFYRHSIDPE.....RHEMHLQTNVYKPGMSFA 1020

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112942 seqs, 26832045 residues

Word size : 1

Total number of hits satisfying chosen parameters: 112925

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC Celerra_SID83/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC Celerra_SID83/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC Celerra_SID83/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC Celerra_SID83/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC Celerra_SID83/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC Celerra_SID83/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC Celerra_SID83/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC Celerra_SID83/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	1.1	763	6	US-10-449-902-43163
2	11	1.1	861	7	US-11-192-437-4
3	9	0.9	94	7	US-11-192-437-13
4	9	0.9	94	7	US-11-192-437-14
5	9	0.9	94	7	US-11-192-437-15
6	9	0.9	471	6	US-10-953-349-19400
7	9	0.9	488	6	US-10-953-349-19399
8	9	0.9	527	6	US-10-953-349-19398
9	8	0.8	108	7	US-11-192-437-16
10	8	0.8	108	7	US-11-192-437-17
11	8	0.8	108	7	US-11-192-437-18
12	8	0.8	108	7	US-11-192-437-19
13	8	0.8	235	6	US-10-539-402-24
14	8	0.8	457	6	US-10-953-349-18475
15	8	0.8	458	6	US-10-953-349-18474
16	8	0.8	501	6	US-10-953-349-18473
17	8	0.8	514	6	US-10-953-349-2683
18	8	0.8	525	6	US-10-953-349-2682
19	8	0.8	547	6	US-10-953-349-2681
20	8	0.8	822	6	US-10-953-349-4066
21	8	0.8	870	6	US-10-953-349-4065
22	8	0.8	876	6	US-10-449-902-37356
23	8	0.8	1011	6	US-10-449-902-56159
24	8	0.8	1050	6	US-10-953-349-4064
25	8	0.8	1052	6	US-10-449-902-53188

26	7	0.7	103	6	US-10-449-902-43508	Sequence 43508, A
27	7	0.7	194	6	US-10-449-902-55668	Sequence 55668, A
28	7	0.7	216	6	US-10-449-902-43156	Sequence 43156, A
29	7	0.7	250	6	US-10-953-349-341	Sequence 341, Appl
30	7	0.7	271	7	US-11-330-363-15	Sequence 15, Appl
31	7	0.7	391	6	US-10-449-902-50154	Sequence 50154, A
32	7	0.7	436	6	US-10-449-902-40680	Sequence 40680, A
33	7	0.7	446	6	US-10-449-902-36916	Sequence 36916, A
34	7	0.7	465	6	US-10-449-902-45325	Sequence 45325, A
35	7	0.7	467	6	US-10-449-902-38097	Sequence 38097, A
36	7	0.7	477	6	US-10-449-902-38723	Sequence 38723, A
37	7	0.7	485	6	US-10-953-349-25267	Sequence 25267, A
38	7	0.7	489	6	US-10-449-902-54513	Sequence 54513, A
39	7	0.7	517	6	US-10-449-902-38630	Sequence 38630, A
40	7	0.7	518	6	US-10-449-902-37736	Sequence 37736, A
41	7	0.7	529	6	US-10-449-902-50447	Sequence 50447, A
42	7	0.7	551	7	US-11-236-238-29	Sequence 29, Appl
43	7	0.7	630	6	US-10-449-902-47718	Sequence 47718, A
44	7	0.7	632	6	US-10-449-902-48865	Sequence 48865, A
45	7	0.7	658	6	US-10-449-902-55772	Sequence 55772, A

ALIGNMENTS

RESULT 1

US-10-449-902-43163
; Sequence 43163, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43163
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-43163

Query Match 1.1%; Score 11; DB 6; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687.

Db 612 LKINAKLGGIN 622

RESULT 2

US-11-192-437-4
; Sequence 4, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO ARGONAUTE PROTEINS

```
;
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-192-437-4
```

```
Query Match 1.1%; Score 11; DB 7; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.016; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;
```

```
QY 677 LKINAKLGGIN 687
Db 557 LKINAKLGGIN 567
|||||
```

```
RESULT 3
US-11-192-437-13
; Sequence 13, Application US/11/192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; TITLE OF INVENTION: ARGONAUTE PROTEINS
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Caenorhabditis briggsae
US-11-192-437-13
```

```
Query Match 0.9%; Score 9; DB 7; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.22; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;
```

```
QY 796 IVVYRDGVS 804
Db 70 IVVYRDGVS 78
|||||
```

```
RESULT 4
US-11-192-437-14
; Sequence 14, Application US/11/192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
```

```
;
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: ARGONAUTE PROTEINS
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-192-437-14
```

```
Query Match 0.9%; Score 9; DB 7; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.22; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;
```

```
QY 796 IVVYRDGVS 804
Db 70 IVVYRDGVS 78
|||||
```

```
RESULT 5
US-11-192-437-15
; Sequence 15, Application US/11/192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; TITLE OF INVENTION: ARGONAUTE PROTEINS
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-192-437-15
```

```
Query Match 0.9%; Score 9; DB 7; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.22; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;
```

```
QY 796 IVVYRDGVS 804
Db 70 IVVYRDGVS 78
|||||
```

```
RESULT 6
US-10-953-349-19400
; Sequence 19400, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
```

```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19400
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-19400

Query Match          0.9%; Score 9; DB 6; Length 471;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGG 685
DB 172 LKINAKLGG 180

RESULT 7
US-10-953-349-19399
; Sequence 19399, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19399
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-19399

Query Match          0.9%; Score 9; DB 6; Length 488;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGG 685
DB 189 LKINAKLGG 197

RESULT 8
US-10-953-349-19398
; Sequence 19398, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19398
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-19398

Query Match          0.9%; Score 9; DB 6; Length 527;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGG 685
DB 189 LKINAKLGG 197
```

```
DB 228 LKINAKLGG 236

RESULT 9
US-11-192-437-16
; Sequence 16, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-192-437-16

Query Match          0.8%; Score 8; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAUVAS 738
DB 16 SIAAUVAS 23

RESULT 10
US-11-192-437-17
; Sequence 17, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-192-437-17

Query Match          0.8%; Score 8; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAUVAS 738
DB 16 SIAAUVAS 23
```

QY 731 SIAAVVAS 738
| | | | | | | |
DB 16 SIAAVVAS 23

RESULT 11
US-11-192-437-18
; Sequence 18, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; TITLE OF INVENTION: ARGONAUTE PROTEINS
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; PRIOR FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Arabidopsis
US-11-192-437-18

Query Match 0.8%; Score 8; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
| | | | | | | |
DB 16 SIAAVVAS 23

RESULT 12
US-11-192-437-19
; Sequence 19, Application US/11192437
; Publication No. US20060141600A1
; GENERAL INFORMATION:
; APPLICANT: Joshua-Tor, Leemor
; APPLICANT: Song, Ji-Joon
; APPLICANT: Hannon, Gregory J.
; APPLICANT: Liu, Jidong
; APPLICANT: Carmell, Michelle A.
; APPLICANT: Rivas, Fabiola
; APPLICANT: Marsden, Carolyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATED TO
; TITLE OF INVENTION: ARGONAUTE PROTEINS
; FILE REFERENCE: CSHL-P01-014
; CURRENT APPLICATION NUMBER: US/11/192,437
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US 60/592,297
; PRIOR FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US 60/592,269
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Arabidopsis zwille
US-11-192-437-19

Query Match 0.8%; Score 8; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
| | | | | | | |
DB 16 SIAAVVAS 23

RESULT 13
US-10-539-402-24
; Sequence 24, Application US/10539402
; Publication No. US20060115477A1
; GENERAL INFORMATION:
; APPLICANT: Xerion Pharmaceuticals AG
; APPLICANT: Tufts University
; TITLE OF INVENTION: Neutropilin-1 Inhibitor
; FILE REFERENCE: XE12EPC
; CURRENT APPLICATION NUMBER: US/10/539,402
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 60/435,893
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: EP 03000615
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 235
; TYPE: PRT
; ORGANISM: human
US-10-539-402-24

Query Match 0.8%; Score 8; DB 6; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 815 ELSRLKSE 822
| | | | | | | |
DB 65 ELSRLKSE 72

RESULT 14
US-10-953-349-18475
; Sequence 18475, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18475
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18475

Query Match 0.8%; Score 8; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
| | | | | | | |
DB 181 SIAAVVAS 188

RESULT 15
US-10-953-349-18474
; Sequence 18474, Application US/109533349
; Publication No. US20060107345A1

; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18474
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-953-349-18474

Query Match 0.8%; Score 8; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
Db 182 SIAAVVAS 189
|||||||

Search completed: July 5, 2006, 14:15:09
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 14:11:13 ; Search time 191 Seconds
(without alignments)
2473.713 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 1020
Sequence: 1 MSNPFLEKGFYRHSLDPE.....RHEMHFLQTNVKYPGMSFA 1020

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2097797 seqs, 463214858 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2096646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	1020	5	US-10-645-746-3
2	1020	100.0	1020	5	US-10-645-735-3
3	1020	100.0	1020	6	US-11-144-985-3
4	818	80.2	818	5	US-10-645-746-13
5	818	80.2	818	5	US-10-645-735-13
6	818	80.2	818	6	US-11-144-985-13
7	11	1.1	820	4	US-10-437-963-131209
8	11	1.1	901	4	US-10-437-963-131210
9	11	1.1	924	4	US-10-408-765A-985
10	10	1.0	530	4	US-10-408-765A-1792
11	10	1.0	580	3	US-09-801-574-16
12	10	1.0	678	4	US-10-425-114-37286
13	9	0.9	117	4	US-10-424-599-244834
14	9	0.9	303	4	US-10-425-115-278723
15	9	0.9	389	4	US-10-174-363-50
16	9	0.9	389	6	US-11-093-888-50
17	9	0.9	500	4	US-10-437-963-138410
18	9	0.9	592	4	US-10-767-701-45656
19	9	0.9	608	4	US-10-437-963-167563
20	9	0.9	700	4	US-10-425-115-302921
21	9	0.9	766	5	US-10-645-746-9
22	9	0.9	766	5	US-10-645-735-9
23	9	0.9	867	4	US-11-144-985-9
24	9	0.9	897	4	US-10-425-115-302918
25	9	0.9	901	4	US-10-174-363-8
26	9	0.9	901	6	US-11-093-888-8
27	9	0.9	904	4	US-10-174-363-55

28	9	0.9	904	6	US-11-093-888-55	Sequence 55, Appl
29	9	0.9	915	4	US-10-174-363-40	Sequence 40, Appl
30	9	0.9	915	6	US-11-093-888-40	Sequence 40, Appl
31	9	0.9	916	4	US-10-174-363-42	Sequence 42, Appl
32	9	0.9	916	4	US-10-424-599-186004	Sequence 186004,
33	9	0.9	916	6	US-11-093-888-42	Sequence 42, Appl
34	9	0.9	930	4	US-10-225-067-2	Sequence 2, Appl
35	9	0.9	930	4	US-10-374-780A-2562	Sequence 2562, Ap
36	9	0.9	2168	4	US-10-437-963-138264	Sequence 138264,
37	8	0.8	104	4	US-10-424-599-187843	Sequence 187843,
38	8	0.8	121	5	US-10-501-282-548	Sequence 548, App
39	8	0.8	161	4	US-10-334-143-192	Sequence 192, App
40	8	0.8	165	4	US-10-767-701-36050	Sequence 36050, A
41	8	0.8	192	4	US-10-424-599-280653	Sequence 280653,
42	8	0.8	192	5	US-10-732-923-6532	Sequence 6532, Ap
43	8	0.8	193	5	US-10-732-923-6531	Sequence 6531, Ap
44	8	0.8	193	5	US-10-732-923-6533	Sequence 6533, Ap
45	8	0.8	209	4	US-10-424-599-182350	Sequence 182350,

ALIGNMENTS

RESULT 1
US-10-645-746-3
; Sequence 3, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabata, Hiroaki
; APPLICANT: Grishok, Allia
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-746-3

Query Match 100.0%; Score 1020; DB 5; Length 1020;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MSNPFLEKGFYRHSLDPEMKWLARPTGKDGKPYEKVLLVNNWPKFSKIYDREYE	60		
Db	1	MSNPFLEKGFYRHSLDPEMKWLARPTGKDGKPYEKVLLVNNWPKFSKIYDREYE	60		
QY	61	YEVKMTKEVLNRKPGKPKKTEIPIDRAKLFQHLRHEKQOTDFTLEDYVDEKDTVY	120		
Db	61	YEVKMTKEVLNRKPGKPKKTEIPIDRAKLFQHLRHEKQOTDFTLEDYVDEKDTVY	120		
QY	121	SVCLRLNTVTSKMLVSEKVKVKKSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE	180		
Db	121	SVCLRLNTVTSKMLVSEKVKVKKSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE	180		
QY	181	EANRSYKFLKNVMTQKRVYAPFVNEBIKQVQAKNFYDNNISILRVPEFSDHPNRFQESLE	240		
Db	181	EANRSYKFLKNVMTQKRVYAPFVNEBIKQVQAKNFYDNNISILRVPEFSDHPNRFQESLE	240		
QY	241	VAPRIENWGIYIGIKELFDGEPVLFNFAVDKLFYNAPKMSILDYLLIIVDQSCNDVVR	300		

```
Db 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLILLIVDQSCNDDVR 300
Qy 301 KDLKTKLMAGKMTIRQAARPRIQALLENILKLCAEVWDMNEMSLRTERHLTFDLCEENSL 360
Db 301 KDLKTKLMAGKMTIRQAARPRIQALLENILKLCAEVWDMNEMSLRTERHLTFDLCEENSL 360
Qy 361 VYKVTGSDRGNAKYDTTLFKIYEENKKFIEFPHLPLVVKVSGAKAYAVPMEHLEVH 420
Db 361 VYKVTGSDRGNAKYDTTLFKIYEENKKFIEFPHLPLVVKVSGAKAYAVPMEHLEVH 420
Qy 421 KPQRYKNRIDLVQMDKFLKRAATRKPHDYKENTLKMKELDPSSEBELNFVERFGLCSKLOM 480
Db 421 KPQRYKNRIDLVQMDKFLKRAATRKPHDYKENTLKMKELDPSSEBELNFVERFGLCSKLOM 480
Qy 481 IECPGKVLKEPMLVNSVNEQIKMTPIVIRGFBQSKQLNVVPEKELCCAFVFNNTAGNPCLE 540
Db 481 IECPGKVLKEPMLVNSVNEQIKMTPIVIRGFBQSKQLNVVPEKELCCAFVFNNTAGNPCLE 540
Qy 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Db 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Qy 601 ATEAKMFERLPDKEQKVLMTIIISKROLNAYGFVKHYCDHTIGVANQHITSETVTTKALA 660
Db 601 ATEAKMFERLPDKEQKVLMTIIISKROLNAYGFVKHYCDHTIGVANQHITSETVTTKALA 660
Qy 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVT 720
Db 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVT 720
Qy 721 HPTSYSGIDYSTAAVVASINPGGTIYRNMIIVTOEBCRPGERAVAHGRERTDILEAKFVKL 780
Db 721 HPTSYSGIDYSTAAVVASINPGGTIYRNMIIVTOEBCRPGERAVAHGRERTDILEAKFVKL 780
Qy 781 LREFAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSEVKQFMSRDEGEDPEPKYT 840
Db 781 LREFAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSEVKQFMSRDEGEDPEPKYT 840
Qy 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
Db 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
Qy 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTMVYDDKGMQODEVYKMTYGLAFLSARC 960
Db 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTMVYDDKGMQODEVYKMTYGLAFLSARC 960

RESULT 2
US-10-645-735-3
; Sequence 3, Application US/10645735
; Publication No. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMG-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-735-3

Query Match      100.0%; Score 1020; DB 5; Length 1020;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSNPPELEKGFYRHSLDPEMKWLARPTGKCDGKFYEKKVLLLVNMFKESSKIYDREYVE 60
Db 1 MSSNPPELEKGFYRHSLDPEMKWLARPTGKCDGKFYEKKVLLLVNMFKESSKIYDREYVE 60
Qy 61 YEYKMTKEVLRNPKPKPPKTEIPIIDRAKLFQWHLRHEKKQTDFILEDYVDEKDTYV 120
Db 61 YEYKMTKEVLRNPKPKPPKTEIPIIDRAKLFQWHLRHEKKQTDFILEDYVDEKDTYV 120
Qy 121 SVCRLLNTVTSKMLVSEKVVKDSKKDEKLEKKILYTMILTYRKFFHLNFSRENPEKDE 180
Db 121 SVCRLLNTVTSKMLVSEKVVKDSKKDEKLEKKILYTMILTYRKFFHLNFSRENPEKDE 180
Qy 181 EANRSYKFLKNVMTQKRYAPFVNEEIKVQFAKNFVYDNNISILRVPESPHDPNRRPEQSL 240
Db 181 EANRSYKFLKNVMTQKRYAPFVNEEIKVQFAKNFVYDNNISILRVPESPHDPNRRPEQSL 240
Qy 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLILLIVDQSCNDDVR 300
Db 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLILLIVDQSCNDDVR 300
Qy 301 KDLKTKLMAGKMTIRQAARPRIQALLENILKLCAEVWDMNEMSLRTERHLTFDLCEENSL 360
Db 301 KDLKTKLMAGKMTIRQAARPRIQALLENILKLCAEVWDMNEMSLRTERHLTFDLCEENSL 360
Qy 361 VYKVTGSDRGNAKYDTTLFKIYEENKKFIEFPHLPLVVKVSGAKAYAVPMEHLEVH 420
Db 361 VYKVTGSDRGNAKYDTTLFKIYEENKKFIEFPHLPLVVKVSGAKAYAVPMEHLEVH 420
Qy 421 KPQRYKNRIDLVQMDKFLKRAATRKPHDYKENTLKMKELDPSSEBELNFVERFGLCSKLOM 480
Db 421 KPQRYKNRIDLVQMDKFLKRAATRKPHDYKENTLKMKELDPSSEBELNFVERFGLCSKLOM 480
Qy 481 IECPGKVLKEPMLVNSVNEQIKMTPIVIRGFBQSKQLNVVPEKELCCAFVFNNTAGNPCLE 540
Db 481 IECPGKVLKEPMLVNSVNEQIKMTPIVIRGFBQSKQLNVVPEKELCCAFVFNNTAGNPCLE 540
Qy 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Db 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Qy 601 ATEAKMFERLPDKEQKVLMTIIISKROLNAYGFVKHYCDHTIGVANQHITSETVTTKALA 660
Db 601 ATEAKMFERLPDKEQKVLMTIIISKROLNAYGFVKHYCDHTIGVANQHITSETVTTKALA 660
Qy 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVT 720
Db 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVT 720
Qy 721 HPTSYSGIDYSTAAVVASINPGGTIYRNMIIVTOEBCRPGERAVAHGRERTDILEAKFVKL 780
Db 721 HPTSYSGIDYSTAAVVASINPGGTIYRNMIIVTOEBCRPGERAVAHGRERTDILEAKFVKL 780
Qy 781 LREFAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSEVKQFMSRDEGEDPEPKYT 840
Db 781 LREFAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSEVKQFMSRDEGEDPEPKYT 840
Qy 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
Db 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
Qy 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTMVYDDKGMQODEVYKMTYGLAFLSARC 960
Db 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTMVYDDKGMQODEVYKMTYGLAFLSARC 960
```

Qy 961 RKPISLPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTHEHEHFLQTNVKGPGMSFA 1020
Db 961 RKPISLPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTHEHEHFLQTNVKGPGMSFA 1020

RESULT 3
US-11-144-985-3
; Sequence 3, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UMY-052CN
; CURRENT APPLICATION NUMBER: US/11/144,985
; PRIOR FILING DATE: 2005-06-03
; PRIOR FILING DATE: 2000-10-13
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

Query Match 100.0%; Score 1020; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNFPLEKGFYRHSIDPEMKMLARTGKDGKFKYKKVLLVNNWPKFSKIYDREYVE 60
Db 1 MSNFPLEKGFYRHSIDPEMKMLARTGKDGKFKYKKVLLVNNWPKFSKIYDREYVE 60

Qy 61 YEVMTEKVLNRKPGKFPKKTETIPIDRAKLFQHLRHEKKTDFILEDYVFEKDTVY 120
Db 61 YEVMTEKVLNRKPGKFPKKTETIPIDRAKLFQHLRHEKKTDFILEDYVFEKDTVY 120

Qy 121 SVCRNLNTVTSKMLVSEKVKVKSDEKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
Db 121 SVCRNLNTVTSKMLVSEKVKVKSDEKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180

Qy 181 EANSYKFLKNVMTQKRYAPFVNEEIKVQFKNFVYDNNNSILRVPSFHDNRFQESLE 240
Db 181 EANSYKFLKNVMTQKRYAPFVNEEIKVQFKNFVYDNNNSILRVPSFHDNRFQESLE 240

Qy 241 VAPRIEAWFGIYIGIKELFGEVPLNFAIVDKLFPYNAKMSLLDYLLIYVDPQSCNDVVR 300
Db 241 VAPRIEAWFGIYIGIKELFGEVPLNFAIVDKLFPYNAKMSLLDYLLIYVDPQSCNDVVR 300

Qy 301 KDLTKLMAKMTIQAARPRIQLLENKLCACAEVNDNEMSLRTERHLTFDLCENSL 360
Db 301 KDLTKLMAKMTIQAARPRIQLLENKLCACAEVNDNEMSLRTERHLTFDLCENSL 360

Qy 361 VYKVTGSDGRNAKYDTTLFKIYENKKFIEFPPLPLVKVKSQAKEYAVPMHEHLEVHE 420
Db 361 VYKVTGSDGRNAKYDTTLFKIYENKKFIEFPPLPLVKVKSQAKEYAVPMHEHLEVHE 420

Qy 421 KPORYKNRIDLVMQDKELKATRKPHDYKENTLKMKELDPSSEELNFRFGLCSKLM 480
Db 421 KPORYKNRIDLVMQDKELKATRKPHDYKENTLKMKELDPSSEELNFRFGLCSKLM 480

Qy 481 IECPGKVLKPMVNSNEQIKMTPTVIRGQEQKLNVPKELCCAVFVNVNAGNFCLE 540
Db 481 IECPGKVLKPMVNSNEQIKMTPTVIRGQEQKLNVPKELCCAVFVNVNAGNFCLE 540

Qy 541 ENDVVVKYTELIGCKEFGIRIGANENRGAQSIMYDATKNYAFYKNCNTLNTGIRPRIA 600
Db 541 ENDVVVKYTELIGCKEFGIRIGANENRGAQSIMYDATKNYAFYKNCNTLNTGIRPRIA 600

Qy 601 ATEAKNMFERLPDKEQKVLMPFIISKQNLNAYGVKHYCDHTIGVANOHIITSETVTKALA 660
Db 601 ATEAKNMFERLPDKEQKVLMPFIISKQNLNAYGVKHYCDHTIGVANOHIITSETVTKALA 660

Qy 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDWSIAISPEEKRRKTMPLTMYVGIDVT 720
Db 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDWSIAISPEEKRRKTMPLTMYVGIDVT 720

Qy 721 HPTSYSIGIDYSIAAVASINPGGTIYRNMIVTQECRPGERBAVAHGRERTDILEAKFVKL 780
Db 721 HPTSYSIGIDYSIAAVASINPGGTIYRNMIVTQECRPGERBAVAHGRERTDILEAKFVKL 780

Qy 781 LRPEAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSEVKQFMSERGEDDEPEPKYT 840
Db 781 LRPEAENNDNRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSEVKQFMSERGEDDEPEPKYT 840

Qy 841 FIVIQKEHNTLLRMEKDKFPVNVNKOLTPAETDVAVAVKQWEEDMKESKETGIWNPSSG 900
Db 841 FIVIQKEHNTLLRMEKDKFPVNVNKOLTPAETDVAVAVKQWEEDMKESKETGIWNPSSG 900

Qy 901 TTVDKLIIVSKYKDFFLASHHGVLTGSRPGHYTVMYDDKQMSQDEVYKQTYGLAPLSARC 960
Db 901 TTVDKLIIVSKYKDFFLASHHGVLTGSRPGHYTVMYDDKQMSQDEVYKQTYGLAPLSARC 960

Qy 961 RKPISLPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTHEHEHFLQTNVKGPGMSFA 1020
Db 961 RKPISLPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTHEHEHFLQTNVKGPGMSFA 1020

RESULT 4
US-10-645-746-13
; Sequence 13, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UMY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; PRIOR FILING DATE: 2003-08-20
; PRIOR FILING DATE: 2000-10-13
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

Query Match 80.2%; Score 818; DB 5; Length 818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 203 VNEEIKVQFAKNFVYDNNNSILRVPSFHDNRFQESLEVAAPRIEAWFGIYIGIKELFGE 262
Db 1 VNEEIKVQFAKNFVYDNNNSILRVPSFHDNRFQESLEVAAPRIEAWFGIYIGIKELFGE 60

Qy 263 PVLNFAIVDKLFPYNAKMSLLDYLLIYVDPQSCNDVVRKDLTKLMAGKMTIQAARPRI 322
Db 61 PVLNFAIVDKLFPYNAKMSLLDYLLIYVDPQSCNDVVRKDLTKLMAGKMTIQAARPRI 120

```
QY 323 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 382
DB 121 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 180
QY 383 KIYENKKEFIEPPHPLPVKVKSGAKAYAVPMHELVHEKPORYKQRIIDLVMQDKFLKRAT 442
DB 181 KIYENKKEFIEPPHPLPVKVKSGAKAYAVPMHELVHEKPORYKQRIIDLVMQDKFLKRAT 240
QY 443 RKPHDYKENTLKMELDFSSSEELNPFVERFGLCSKLQMIETCPGKVLKEPMLVNSVNEQIK 502
DB 241 RKPHDYKENTLKMELDFSSSEELNPFVERFGLCSKLQMIETCPGKVLKEPMLVNSVNEQIK 300
QY 503 MTPVIRGFOEKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKPFGIRI 562
DB 301 MTPVIRGFOEKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKPFGIRI 360
QY 563 GANENRGAQSIWYDATKNEYAFYKNCNTLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 622
DB 361 GANENRGAQSIWYDATKNEYAFYKNCNTLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 420
QY 623 IISKQLNAYGFVKHYCDHTIGVANOHTITSETVTKALASLRHEKSGKRIFYQIALKINAK 682
DB 421 IISKQLNAYGFVKHYCDHTIGVANOHTITSETVTKALASLRHEKSGKRIFYQIALKINAK 480
QY 683 LGGINQELDWSIAETISPEEKERRKTMPLTMVYVIGIDVTHPTSYSGIDYSIAAVVASINPG 742
DB 481 LGGINQELDWSIAETISPEEKERRKTMPLTMVYVIGIDVTHPTSYSGIDYSIAAVVASINPG 540
QY 743 GTIYRNMIVTQECRPGERAVAHGERTDILEAKFVKLLRREFAENNDRAPAHIVVYRDG 802
DB 541 GTIYRNMIVTQECRPGERAVAHGERTDILEAKFVKLLRREFAENNDRAPAHIVVYRDG 600
QY 803 VSDSEMLRVSHDELRSLSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 862
DB 601 VSDSEMLRVSHDELRSLSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 660
QY 863 VNKDLPARTDVAVAVKQWEDMKESKETGIVNPSGGTTVDKLVSKYKFPFFLASHHG 922
DB 661 VNKDLPARTDVAVAVKQWEDMKESKETGIVNPSGGTTVDKLVSKYKFPFFLASHHG 720
QY 923 VLGTSRPGHYTWMYDDKQMSQDEVYKMTYGLAFLSARCKPISLPVPVHYAHLSCCKAKE 982
DB 721 VLGTSRPGHYTWMYDDKQMSQDEVYKMTYGLAFLSARCKPISLPVPVHYAHLSCCKAKE 780
QY 983 LYRTYKEHYIGDYAOPRTRHEMEHFLQTNVKYPGMSFA 1020
DB 781 LYRTYKEHYIGDYAOPRTRHEMEHFLQTNVKYPGMSFA 818
RESULT 5
US-10-645-735-13
; Sequence 13, Application US/10645735
; Publication No. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMG-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
```

```
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-735-13
```

```
Query Match 80.2%; Score 818; DB 5; Length 818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 203 VNEEIKVQFAKNFVYDNNISILRVPSFHDPNRFEQSLVAPRIEAWFGIYIGIKELFDGE 262
DB 1 VNEEIKVQFAKNFVYDNNISILRVPSFHDPNRFEQSLVAPRIEAWFGIYIGIKELFDGE 60
QY 263 PVLNFAIVDKLPYNAPKMSLLDYLLIVDPQSCNDNDVRKDLATKLMAGKWTIRQAAARPI 322
DB 61 PVLNFAIVDKLPYNAPKMSLLDYLLIVDPQSCNDNDVRKDLATKLMAGKWTIRQAAARPI 120
QY 323 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 382
DB 121 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 180
QY 383 KIYENKKEFIEPPHPLPVKVKSGAKAYAVPMHELVHEKPORYKQRIIDLVMQDKFLKRAT 442
DB 181 KIYENKKEFIEPPHPLPVKVKSGAKAYAVPMHELVHEKPORYKQRIIDLVMQDKFLKRAT 240
QY 443 RKPHDYKENTLKMELDFSSSEELNPFVERFGLCSKLQMIETCPGKVLKEPMLVNSVNEQIK 502
DB 241 RKPHDYKENTLKMELDFSSSEELNPFVERFGLCSKLQMIETCPGKVLKEPMLVNSVNEQIK 300
QY 503 MTPVIRGFOEKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKPFGIRI 562
DB 301 MTPVIRGFOEKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKPFGIRI 360
QY 563 GANENRGAQSIWYDATKNEYAFYKNCNTLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 622
DB 361 GANENRGAQSIWYDATKNEYAFYKNCNTLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 420
QY 623 IISKQLNAYGFVKHYCDHTIGVANOHTITSETVTKALASLRHEKSGKRIFYQIALKINAK 682
DB 421 IISKQLNAYGFVKHYCDHTIGVANOHTITSETVTKALASLRHEKSGKRIFYQIALKINAK 480
QY 683 LGGINQELDWSIAETISPEEKERRKTMPLTMVYVIGIDVTHPTSYSGIDYSIAAVVASINPG 742
DB 481 LGGINQELDWSIAETISPEEKERRKTMPLTMVYVIGIDVTHPTSYSGIDYSIAAVVASINPG 540
QY 743 GTIYRNMIVTQECRPGERAVAHGERTDILEAKFVKLLRREFAENNDRAPAHIVVYRDG 802
DB 541 GTIYRNMIVTQECRPGERAVAHGERTDILEAKFVKLLRREFAENNDRAPAHIVVYRDG 600
QY 803 VSDSEMLRVSHDELRSLSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 862
DB 601 VSDSEMLRVSHDELRSLSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 660
QY 863 VNKDLPARTDVAVAVKQWEDMKESKETGIVNPSGGTTVDKLVSKYKFPFFLASHHG 922
DB 661 VNKDLPARTDVAVAVKQWEDMKESKETGIVNPSGGTTVDKLVSKYKFPFFLASHHG 720
QY 923 VLGTSRPGHYTWMYDDKQMSQDEVYKMTYGLAFLSARCKPISLPVPVHYAHLSCCKAKE 982
DB 721 VLGTSRPGHYTWMYDDKQMSQDEVYKMTYGLAFLSARCKPISLPVPVHYAHLSCCKAKE 780
QY 983 LYRTYKEHYIGDYAOPRTRHEMEHFLQTNVKYPGMSFA 1020
DB 781 LYRTYKEHYIGDYAOPRTRHEMEHFLQTNVKYPGMSFA 818
```

RESULT 6

```
US-11-144-985-13
; Sequence 13, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
```

```
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMY-052CN
; CURRENT APPLICATION NUMBER: US/11/144,985
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-144-985-13

Query Match      80.2%; Score 818; DB 6; Length 818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 VNEEIKVQFAKFNVDNNSILRVPESPHDNRPEQSLEVPRIEAWFGIYIGIKELPDGE 262
DB 1 VNEEIKVQFAKFNVDNNSILRVPESPHDNRPEQSLEVPRIEAWFGIYIGIKELPDGE 60

QY 263 PVLNFAIVDKLFLYNAPKMSLDYLLLVDPQSCNDDYRKDKTKLMAGKMTIRQAAPRI 322
DB 61 PVLNFAIVDKLFLYNAPKMSLDYLLLVDPQSCNDDYRKDKTKLMAGKMTIRQAAPRI 120

QY 323 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 382
DB 121 ROLLENLKLKCAEVDNEMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNAKKYDTTLF 180

QY 383 KIYEENKFTFPHLPLVVKYSGAKYAVPMHEHLEHVEKPORYKNRIDLVQDKFLKRA 442
DB 181 KIYEENKFTFPHLPLVVKYSGAKYAVPMHEHLEHVEKPORYKNRIDLVQDKFLKRA 240

QY 443 RKPHDYKENTLKMELKELDPSSBELNVERFGLCKLQMIKPCGKVLKEPMLVNSVNEQIK 502
DB 241 RKPHDYKENTLKMELKELDPSSBELNVERFGLCKLQMIKPCGKVLKEPMLVNSVNEQIK 300

QY 503 MTPVIRGFQSKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKFRGIRI 562
DB 301 MTPVIRGFQSKQLNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGCKFRGIRI 360

QY 563 GANENRGAQSIMYDQNEVAFYKNCYTLNTGIGRFEIAATEAKNMFERLPDKSQKVLMPFI 622
DB 361 GANENRGAQSIMYDQNEVAFYKNCYTLNTGIGRFEIAATEAKNMFERLPDKSQKVLMPFI 420

QY 623 IISKROLNAGFVKHYCDHTIGVANOHITSETVTKALASIRHEKSGKRIFYQIALKINAK 682
DB 421 IISKROLNAGFVKHYCDHTIGVANOHITSETVTKALASIRHEKSGKRIFYQIALKINAK 480

QY 683 LGGINQELQWSEIAEISPEEKERRKTMPLTMYVIGIDVTHPTSYSGIDYSIAAASINPG 742
DB 481 LGGINQELQWSEIAEISPEEKERRKTMPLTMYVIGIDVTHPTSYSGIDYSIAAASINPG 540

QY 743 GTTYRNMIVTQESCRPGERAVAGRERTDILEAKFVKLLREFANNDRNAPAHIVYVRDG 802
DB 541 GTTYRNMIVTQESCRPGERAVAGRERTDILEAKFVKLLREFANNDRNAPAHIVYVRDG 600

QY 803 VSDSEMLRVSHDELRSIKSVKOPMSERGEDPEPKYTFIVIOKRNTRLIRMEWKQPV 862
DB 601 VSDSEMLRVSHDELRSIKSVKOPMSERGEDPEPKYTFIVIOKRNTRLIRMEWKQPV 660

QY 863 VNKDLTPAETDVAVAQKQWEEDMKESKETGIWNPSSGTTVDKLIIVSKYKFDFFLASHHG 922
DB 661 VNKDLTPAETDVAVAQKQWEEDMKESKETGIWNPSSGTTVDKLIIVSKYKFDFFLASHHG 720
```

```
QY 923 VLGTSRPGHYTVMYDDKMGSDQEVYKMTYGLAFLSARCRKPSLPPVPHYAHLSCCKAKE 982
DB 721 VLGTSRPGHYTVMYDDKMGSDQEVYKMTYGLAFLSARCRKPSLPPVPHYAHLSCCKAKE 780

QY 983 LYRTYKEHYIGDYAQPTRHEMEHFLQTNVYKPGMSFA 1020
DB 781 LYRTYKEHYIGDYAQPTRHEMEHFLQTNVYKPGMSFA 818
```

RESULT 7

```
US-10-437-963-131209
; Sequence 131209, Application US/10437963
; Publication No. US20040123343A1
```

; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131209
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33297C.1.pep
US-10-437-963-131209
```

```
Query Match      1.1%; Score 11; DB 4; Length 820;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 677 LKINAKLGGIN 687
DB 574 LKINAKLGGIN 584
```

RESULT 8

```
US-10-437-963-131210
; Sequence 131210, Application US/10437963
; Publication No. US20040123343A1
```

; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131210
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: (1)..(901)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
```

; OTHER INFORMATION: Clone ID: PAT_MRT4530_33298C.1.pap
US-10-437-963-131210

Query Match 1.1%; Score 11; DB 4; Length 901;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
Db 602 LKINAKLGGIN 612

RESULT 9
US-10-408-765A-995
; Sequence 995, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 995
; LENGTH: 924
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-995

Query Match 1.1%; Score 11; DB 4; Length 924;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
Db 620 LKINAKLGGIN 630

RESULT 10
US-10-408-765A-1792
; Sequence 1792, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1792
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1792

Query Match 1.0%; Score 10; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
Db 367 IVVYRDGVSD 376

RESULT 11
US-09-801-574-16
; Sequence 16, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijing Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-574-16

Query Match 1.0%; Score 10; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
Db 417 IVVYRDGVSD 426

RESULT 12
US-10-425-114-37286
; Sequence 37286, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37286
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB83-005-F5_FLI.pap
US-10-425-114-37286

Query Match 1.0%; Score 10; DB 4; Length 678;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 STAAVVASIN 740
Db 440 STAAVVASIN 449

RESULT 13

US-10-424-599-244834
; Sequence 244834, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 244834
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63114C.1.pep
US-10-424-599-244834

Query Match 0.9%; Score 9; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 677 LKINAKLGG 685
|||||||
Db 60 LKINAKLGG 68

RESULT 14

US-10-425-115-278723
; Sequence 278723, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 278723
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(303)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_185770C.1.pep
US-10-425-115-278723

Query Match 0.9%; Score 9; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 10; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 677 LKINAKLGG 685
|||||||
Db 5 LKINAKLGG 13

RESULT 15

US-10-174-363-50
; Sequence 50, Application US/10174363
; Publication No. US20030077623A1
; GENERAL INFORMATION:

; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptional
; TITLE OF INVENTION: Gene Silencing
; FILE REFERENCE: BB1454 US NA
; CURRENT APPLICATION NUMBER: US/10/174,363
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,973
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 50
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-174-363-50

Query Match 0.9%; Score 9; DB 4; Length 389;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 677 LKINAKLGG 685
|||||||
Db 89 LKINAKLGG 97

Search completed: July 5, 2006, 14:14:42
Job time : 193 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:59:07 ; Search time 52 Seconds
(without alignments)
1716.947 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 1020

Sequence: 1 MSSNPFLEKGFVHSLDPE.....RHEMFLOTNVYKPGMSFA 1020

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 650591 seqs, 87530628 residues

Word size : 1

Total number of hits satisfying chosen parameters: 649417

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Celerra_SID33/ptodata/2/iaa/5 COMB.pep:*
- 2: /EMC Celerra_SID33/ptodata/2/iaa/6 COMB.pep:*
- 3: /EMC Celerra_SID33/ptodata/2/iaa/7 COMB.pep:*
- 4: /EMC Celerra_SID33/ptodata/2/iaa/H COMB.pep:*
- 5: /EMC Celerra_SID33/ptodata/2/iaa/PCTUS COMB.pep:*
- 6: /EMC Celerra_SID33/ptodata/2/iaa/RE COMB.pep:*
- 7: /EMC Celerra_SID33/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	0.8	179	2	US-09-270-767-40274
2	8	0.8	179	2	US-09-270-767-55490
3	8	0.8	243	1	US-08-685-992-35
4	8	0.8	243	1	US-09-144-925-35
5	8	0.8	445	2	US-09-328-352-6669
6	8	0.8	452	2	US-10-332-795-7
7	8	0.8	969	2	US-09-533-029-106
8	7	0.7	15	1	US-08-268-251-19
9	7	0.7	15	5	PCT-US93-01112-19
10	7	0.7	23	1	US-08-086-428B-119
11	7	0.7	23	1	US-08-268-251-49
12	7	0.7	23	1	US-08-468-570-159
13	7	0.7	23	1	US-08-290-665A-263
14	7	0.7	23	2	US-08-466-601A-159
15	7	0.7	23	5	PCT-US93-01112-43
16	7	0.7	23	5	PCT-US95-10398-263
17	7	0.7	67	1	US-09-248-796A-25158
18	7	0.7	77	1	US-08-774-065-6
19	7	0.7	80	2	US-09-270-767-56655
20	7	0.7	97	2	US-09-621-976-5201
21	7	0.7	110	2	US-09-513-599C-4110
22	7	0.7	120	2	US-09-219-983A-23
23	7	0.7	120	2	US-10-114-774-23
24	7	0.7	137	2	US-09-913-204-17
25	7	0.7	162	2	US-09-270-767-41434
26	7	0.7	163	2	US-09-270-767-36487

27 7 0.7 163 2 US-09-270-767-51704 Sequence 51704, A
28 7 0.7 175 2 US-09-248-796A-14297 Sequence 14297, A
29 7 0.7 186 2 US-09-270-767-46614 Sequence 46614, A
30 7 0.7 187 2 US-09-248-796A-15309 Sequence 15309, A
31 7 0.7 192 1 US-08-086-428B-102 Sequence 102, App
32 7 0.7 192 1 US-08-468-570-102 Sequence 102, App
33 7 0.7 192 1 US-08-290-665A-102 Sequence 102, App
34 7 0.7 192 2 US-08-466-601A-102 Sequence 102, App
35 7 0.7 192 5 PCT-US95-10398-102 Sequence 102, App
36 7 0.7 198 1 US-08-682-517-19 Sequence 19, Appl
37 7 0.7 217 2 US-09-543-681A-7649 Sequence 7649, Ap
38 7 0.7 217 2 US-09-543-681A-7862 Sequence 7862, Ap
39 7 0.7 219 1 US-08-208-885-2 Sequence 2, Appli
40 7 0.7 219 1 US-08-462-177-2 Sequence 2, Appli
41 7 0.7 219 1 US-08-833-622-2 Sequence 2, Appli
42 7 0.7 220 2 US-09-270-767-38762 Sequence 38762, A
43 7 0.7 220 2 US-09-270-767-53979 Sequence 53979, A
44 7 0.7 223 2 US-09-248-796A-27564 Sequence 27564, A
45 7 0.7 241 2 US-09-134-000C-3623 Sequence 3623, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-40274
; Sequence 40274, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40274
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-40274

Query Match 0.8%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
|||
Db 26 SIAAVVAS 33

RESULT 2
US-09-270-767-55490
; Sequence 55490, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55490
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55490

Query Match 0.8%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVAS 738
DB 26 SIAAVVAS 33

RESULT 3

US-08-685-992-35
; Sequence 35, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992
; FILING DATE: 25-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-03
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:

INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-685-992-35

Query Match 0.8%; Score 8; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 VTKALASL 662
DB 149 VTKALASL 156

RESULT 4

US-09-144-925-35
; Sequence 35, Application US/09144925
; Patent No. 5951979
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES

NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02421-4799

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,925
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/685,992
; FILING DATE: July 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-03Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:

INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-144-925-35

Query Match 0.8%; Score 8; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 VTKALASL 662
DB 149 VTKALASL 156

RESULT 5

US-09-328-352-6669
; Sequence 6669, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6669
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6669

Query Match 0.8%; Score 8; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 SLLDYLL 288
DB 97 SLLDYLL 104

RESULT 6
US-10-332-795-7
; Sequence 7, Application US/10332795
; Patent No. 6946253
; GENERAL INFORMATION:
; APPLICANT: The University Court of the University of Glasgow
; APPLICANT: Clements, John Barlie
; APPLICANT: MacLean, Alasdair Roderick
; TITLE OF INVENTION: HERPES ZINC FINGER MOTIFS
; FILE REFERENCE: 9013-51
; CURRENT APPLICATION NUMBER: US/10/332.795
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: GB 0016890.6
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Varicella-zoster virus
US-10-332-795-7

Query Match 0.8%; Score 8; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 LLENLKLK 332
|||||
DB 321 LLENLKLK 328

RESULT 7
US-09-533-029-106
; Sequence 106, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 106
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1149
US-09-533-029-106

Query Match 0.8%; Score 8; DB 2; Length 969;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 STAAVVAS 738
|||||
DB 706 STAAVVAS 713

RESULT 8
US-08-268-251-19
; Sequence 19, Application US/08268251
; Patent No. 5585475
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,251
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-268-251-19

Query Match 0.7%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 815 ELRLSKS 821
|||||
DB 6 ELRLSKS 12

RESULT 9
PCT-US93-01112-19
; Sequence 19, Application PC/TUS9301112
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01112
FILING DATE: 19930208
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 272.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-01112-19

Query Match 0.7%; Score 7; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 815 ELRSLKS 821
DB 6 ELRSLKS 12

RESULT 10
US-08-086-428B-159
Sequence 159, Application US/08086428B
Patent No. 5514539
GENERAL INFORMATION:
APPLICANT: BUGH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,428B
FILING DATE: 29-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: 421792
INFORMATION FOR SEQ ID NO: 159:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-086-428B-159

Query Match 0.7%; Score 7; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 SILRVPE 227
DB 8 SILRVPE 14

RESULT 11
US-08-268-251-49
Sequence 49, Application US/08268251
Patent No. 5585475
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A
APPLICANT: Dedman, John R
APPLICANT: Kaetzel, Marcia A
TITLE OF INVENTION: Calmodulin-Binding Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,251
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 272.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-268-251-49

Query Match 0.7%; Score 7; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 815 ELRSLKS 821
DB 8 ELRSLKS 14

RESULT 12
US-08-468-570-159
Sequence 159, Application US/08468570

Patent No. 5871962
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,570
FILING DATE: 6-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-468-570-159
Query Match 0.7%; Score 7; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 221 SILRVPE 227
Db 8 SILRVPE 14
RESULT 13
US-08-290-665A-263
Sequence 263, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK

COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 263:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-290-665A-263
Query Match 0.7%; Score 7; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 221 SILRVPE 227
Db 8 SILRVPE 14
RESULT 14
US-08-466-601A-159
Sequence 159, Application US/08466601A
Patent No. 6572864
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,601A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US2

Search completed: July 5, 2006, 14:00:40
Job time : 54 secs

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-466-601A-159

Query Match 0.7%; Score 7; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 SILRVPE 227
Db 8 SILRVPE 14

RESULT 15
PCT-US93-01112-49
Sequence 49, Application PC/TUS9301112
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A
APPLICANT: Dedman, John R
TITLE OF INVENTION: Calmodulin-Binding Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01112
FILING DATE: 19930208
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 272.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-01112-49

Query Match 0.7%; Score 7; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 815 ELRLSKS 821
Db 8 ELRLSKS 14

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:50:21 ; Search time 316 Seconds
(without alignments)
2985.810 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 1020

Sequence: 1 MSSNPPELEKGFYRHSILDPE.....RHEMFLOQTNVYKPGMSFA 1020

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Uniprot 7.2.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	1020	2	Q9XU82 CAEL
2	26	2.5	976	2	Q60V80 CAER
3	15	1.5	185	2	Q46002 CAEL
4	11	1.1	387	2	Q9NXV9 HUMAN
5	11	1.1	794	2	Q5ZMW0 CHICK
6	11	1.1	861	1	I2C4 HUMAN
7	11	1.1	861	2	Q5VXF0 HUMAN
8	11	1.1	884	2	Q4KLVE XENLA
9	11	1.1	945	2	Q4SVE6 TETNG
10	10	1.0	371	2	Q9NW28 HUMAN
11	10	1.0	455	2	Q3TQ88 MOUSE
12	10	1.0	530	2	Q96SW6 HUMAN
13	10	1.0	580	2	Q99MV6 MOUSE
14	10	1.0	971	2	Q8CDG1 MOUSE
15	10	1.0	971	2	Q9JMB6 MOUSE
16	10	1.0	973	2	Q8TCS9 HUMAN
17	10	1.0	1054	2	Q75HC2 ORYSA
18	10	1.0	1121	2	Q61931 CAEL
19	9	0.9	361	2	Q7XBH5 ORYSA
20	9	0.9	414	2	Q9KRL3 VIBCH
21	9	0.9	478	2	Q5ZHV0 CHICK
22	9	0.9	483	2	Q31Q93 SYNFP
23	9	0.9	483	2	Q5N323 SYNFP
24	9	0.9	540	2	Q3GBT1 9FTRM
25	9	0.9	580	2	Q56X15 ARATH
26	9	0.9	580	2	Q358Y1 9BRAD
27	9	0.9	850	2	Q3E984 ARATH
28	9	0.9	879	2	Q5NEN9 ORYSA
29	9	0.9	889	2	Q6VJ55 ORYSA
30	9	0.9	891	2	Q86B39 CAEL
31	9	0.9	892	2	Q84VQ0 ARATH

32 9 0.9 896 2 Q84Y14 ARATH
33 9 0.9 904 2 Q9SDG8 ORYSA
34 9 0.9 905 2 Q2LFC1 NICBE
35 9 0.9 910 2 Q16720 CAEL
36 9 0.9 912 2 Q2LFC2 NICBE
37 9 0.9 924 2 Q9ZVD5 ARATH
38 9 0.9 948 2 Q4RPA4 TETNG
39 9 0.9 1002 2 Q20578 CAEL
40 9 0.9 1010 2 Q3LTR7 CAEL
41 9 0.9 1024 2 Q61PV1 CAEL
42 9 0.9 1032 2 Q7JLZ2 CAEL
43 9 0.9 1034 2 Q61J44 CAER
44 9 0.9 1035 2 Q21079 CAEL
45 9 0.9 1040 1 Y043 CAEL

ALIGNMENTS

RESULT 1
Q9XU82 CAEL
ID Q9XU82 CAEL PRELIMINARY; PRT; 1020 AA.
AC Q9XU82; Q9UGQ1;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2001, sequence version 2.
DT 07-FEB-2006, entry version 26.
DE Hypothetical protein rde-1 (RNA interference promoting factor RDE-1).
GN Name=rde-1; ORFNames=K08H10.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99059613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20004389; PubMed=10535731; DOI=10.1016/S0092-8674(00)81644-X;
RA Tabara H., Sarkissian M., Kelly W.G., Pleener J., Grishok A., Timmons L., Fire A., Mello C.C.;
RT "The rde-1 gene, RNA interference, and transposon silencing in C. elegans."
RL Cell 99:123-132(1999).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
EMBL: Z83113; CAB05546.2; -; Genomic DNA.
EMBL: AF180730; AAP06159.1; -; mRNA.
PIR: T23510; T23510.
DR Ensembl; K08H10.7; Caenorhabditis elegans.
DR WormBase; WBGene0004323; rde-1.
DR WormPep; K08H10.7; CS28243.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS0822; PIWI; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1020 AA; 118804 MW; 8E2F1A2BFC43A670 CRC64;

Query Match 100.0%; Score 1020; DB 2; Length 1020;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNPPELEKGFYRHSILDPEMKWLARPTCKDCKGKFKYKKVLLVNNFKSSKLYDREY 60
DB 1 MSSNPPELEKGFYRHSILDPEMKWLARPTCKDCKGKFKYKKVLLVNNFKSSKLYDREY 60

```
QY 61 YEYKMTKEVLRKPGKPPKKTETIPIDRAKLFWOHLRHEKKQTDPILDYVPDEKDTY 120
DB 61 YEYKMTKEVLRKPGKPPKKTETIPIDRAKLFWOHLRHEKKQTDPILDYVPDEKDTY 120
QY 121 SVCLRLNTVTSKMLVSKVKKDEKKDEKLEKKILYTMILTYRKKFPHLNFSGREPEKDE 180
DB 121 SVCLRLNTVTSKMLVSKVKKDEKKDEKLEKKILYTMILTYRKKFPHLNFSGREPEKDE 180
QY 181 EANRSYKFLKNVMTQKRVYAPFVNEBIKQVAKNFVYDNNSTLRVPESPHDNNRPFQSL 240
DB 181 EANRSYKFLKNVMTQKRVYAPFVNEBIKQVAKNFVYDNNSTLRVPESPHDNNRPFQSL 240
QY 241 VAPRIEAWGIYIGIKELFGEPEVLNFAVDKLFYNAPKMSLLDYLLIIVDQSCNDDYR 300
DB 241 VAPRIEAWGIYIGIKELFGEPEVLNFAVDKLFYNAPKMSLLDYLLIIVDQSCNDDYR 300
QY 301 KDLKTKLMAGKMTIROAARPRIQLLENLKLCAEVDNEMSLRTERHLTFDLDCENSL 360
DB 301 KDLKTKLMAGKMTIROAARPRIQLLENLKLCAEVDNEMSLRTERHLTFDLDCENSL 360
QY 361 VYKVTGSDRGNAKKYDTTLFKIYEENKKFIEFPHLPLVKVKGAKYAVPMHEHLEVHE 420
DB 361 VYKVTGSDRGNAKKYDTTLFKIYEENKKFIEFPHLPLVKVKGAKYAVPMHEHLEVHE 420
QY 421 KPORYKNRIDLVMQDFLKRATKPHDYKENTLKMKELDFFSSEELNFVERGLCSKLOM 480
DB 421 KPORYKNRIDLVMQDFLKRATKPHDYKENTLKMKELDFFSSEELNFVERGLCSKLOM 480
QY 481 IECPCGVKLEPMLVNSVQIKMTPIRVIRFOEQKQLNVPEKELCCAFVFNNTAGNPCLE 540
DB 481 IECPCGVKLEPMLVNSVQIKMTPIRVIRFOEQKQLNVPEKELCCAFVFNNTAGNPCLE 540
QY 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSTMYDATKNEYAFYKCTLNTGIGRFETA 600
DB 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSTMYDATKNEYAFYKCTLNTGIGRFETA 600
QY 601 ATEAKNMFRLPDKEOKVLMFIIISKQLNAYGFVKHYCDHTIGVANOHTSETVTTKALA 660
DB 601 ATEAKNMFRLPDKEOKVLMFIIISKQLNAYGFVKHYCDHTIGVANOHTSETVTTKALA 660
QY 661 SLRHEKSKRIFQYIALKINAKLGINQBELDSEIAEISPEEKERRKTMPLTYMGIDVT 720
DB 661 SLRHEKSKRIFQYIALKINAKLGINQBELDSEIAEISPEEKERRKTMPLTYMGIDVT 720
QY 721 HPTSYSGIDYSIAAVVASINPGGTIYRNMIVTQEECPGERAVAHGRERTDILEAKFVKL 780
DB 721 HPTSYSGIDYSIAAVVASINPGGTIYRNMIVTQEECPGERAVAHGRERTDILEAKFVKL 780
QY 781 LREFAENNDNRPAHIVVYRDGVSDSEMLRVSHDELSLKSEVKQFMSRDEGDEPKYT 840
DB 781 LREFAENNDNRPAHIVVYRDGVSDSEMLRVSHDELSLKSEVKQFMSRDEGDEPKYT 840
QY 841 FVIOQRHNTLRLRMKDKPVNKLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
DB 841 FVIOQRHNTLRLRMKDKPVNKLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
QY 901 TTVDKLIIVSKYKPDFFLASHHGVLTGSRPGHYTVMYDDKGMQDEYVKMTYGLAFLSARC 960
DB 901 TTVDKLIIVSKYKPDFFLASHHGVLTGSRPGHYTVMYDDKGMQDEYVKMTYGLAFLSARC 960
QY 961 RKPISLIPVPHVAHLSCAKELYRTYKEHYIGDYAQPRTHEMEHFLQTNVKYPCWMSFA 1020
DB 961 RKPISLIPVPHVAHLSCAKELYRTYKEHYIGDYAQPRTHEMEHFLQTNVKYPCWMSFA 1020

RESULT 2
Q60VS0_CAEBR
ID Q60VS0_CAEBR PRELIMINARY; PRT; 976 AA.
AC Q60VS0;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
```

```
DE Hypothetical protein CBGI9426.
GN Name=CBGI9426;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]_TaxID=6238;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF16;
RX PubMed=14624247; DOI=10.1371/journal.pbio.0000045;
RA Stein L.D., Bao Z., Blasgar D., Blumenthal T., Brent M.R., Chen N.,
RA Chinwalla A., Clarke L., Clee C., Coghlan A., Coulson A.,
RA D'Eustachio P., Fitch D.H.A., Fulton L.A., Fulton R.E.,
RA Griffiths-Jones S., Harris T.W., Hillier L.W., Kamath R.,
RA Kuwabara P.E., Plumb E.R., Marra M.A., Miner T.L., Mink P.,
RA Mullikin J.C., Plumb R.W., Rogers J., Schein J.E., Sohrmann M.,
RA Spieth J., Stajich J.E., Wei C., Willey D., Wilson R.K., Durbin R.,
RA Waterston R.H.;
RT "The genome sequence of Caenorhabditis briggsae: a platform for
RT comparative genomics";
RL PLoS Biol. 1:166-192(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; CAAC0100098; CAB72296.1; -; Genomic_DNA.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50822; PIWI; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 976 AA; 114611 MW; 8B5427451C9ADCAA CRC64;

Query Match 2.5%; Score 26; DB 2; Length 976;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 HIVVYRDGVSDSEMLRVSHDELSLK 820
DB 763 HIVVYRDGVSDSEMLRVSHDELSLK 788

RESULT 3
O46002 CAEBL PRELIMINARY; PRT; 185 AA.
AC O46002;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE Hypothetical protein.
GN ORFNames=ZK218.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_TaxID=6239;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RT "Genome sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; Z82085; CAB04988.1; -; Genomic_DNA.
DR PIR; T27784; T27784.
```



```
DR Ensembl; ZK218.8; Caenorhabditis elegans.
DR WormBase; WBGene00013942; ZK218.8.
DR WormPep; ZK218.8; CE16708.
DR InterPro; IPR003165; Pfam.
DR Pfam; PF02171; Pfam; 1.
DR PROSITE; PS50822; Pfam; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 185 AA; 21081 MW; B79F7862584F11A5 CRC64;
SQ SEQUENCE 185 AA; 21081 MW; B79F7862584F11A5 CRC64;

Query Match 1.5%; Score 15; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 LAFLSARCRKPSLIP 967
DB 156 LAFLSARCRKPSLIP 170

RESULT 4
Q9NXV9_HUMAN PRELIMINARY; PRT; 387 AA.
AC Q9NXV9;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Hypothetical protein FLJ20033.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DE EMBL; AK000040; BAA90899.1; -; mRNA.
DR Ensembl; ENSG00000134698; Homo sapiens.
DR InterPro; IPR003165; Pfam.
DR Pfam; PF02171; Pfam; 1.
DR PROSITE; PS50822; Pfam; 1.
SQ SEQUENCE 387 AA; 43196 MW; 38ABC7BDF0A585B0 CRC64;

Query Match 1.1%; Score 11; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
DB 83 LKINAKLGGIN 93

RESULT 5
Q5ZMW0_CHICK PRELIMINARY; PRT; 794 AA.
AC Q5ZMW0;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Hypothetical protein.
GN ORFNames=RCJMB04.1a17;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DE EMBL; AK000040; BAA90899.1; -; mRNA.
DR Ensembl; ENSG00000134698; Homo sapiens.
DR InterPro; IPR003165; Pfam.
DR Pfam; PF02171; Pfam; 1.
DR PROSITE; PS50822; Pfam; 1.
SQ SEQUENCE 387 AA; 43196 MW; 38ABC7BDF0A585B0 CRC64;

Query Match 1.1%; Score 11; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
DB 83 LKINAKLGGIN 93

RESULT 6
I2C4_HUMAN STANDARD; PRT; 861 AA.
AC Q9HCK5;
DT 14-NOV-2003, integrated into UniProtKB/Swiss-Prot.
DT 14-NOV-2003, sequence version 2.
DT 07-FEB-2006, entry version 21.
DE Eukaryotic translation initiation factor 2C 4 (eIF2C 4) (eIF-2C 4)
DE (Argonaute-4).
GN Name=EIF2C4; Synonyms=AGO4, KIAA1567;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
MEDLINE=20450683; PubMed=10997877; DOI=10.1093/dnares/7.4.271;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
CC -!- FUNCTION: Plays an important role in the eukaryotic peptide chain
CC initiation process (By similarity).
CC -!- SIMILARITY: Belongs to the argonaute family.
CC -!- SIMILARITY: Contains 1 PAZ domain.
CC -!- SIMILARITY: Contains 1 Piwi domain.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DE EMBL; AB046787; BAB13393.1; ALT_INIT; mRNA.
DR SMR; Q9HCK5; 212-374.
DR Ensembl; ENSG00000134698; Homo sapiens.
DR HGNC; HGNC:18424; EIF2C4.
DR MIM; 607356; Gene.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003185; Piwi.
DR Pfam; PF02170; PAZ; 1.
```

DR Pfam; PF02171; Pfam; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
KW Initiation factor; Protein biosynthesis.
FT CHAIN 1 861
FT
FT
FT DOMAIN 225 338 /FTID=PRO_0000194063.
FT DOMAIN 509 820 PAZ.
FT SEQUENCE 861 AA; 97097 MW; F236FF05047534C1 CRC64;
SQ
Query Match 1.1%; Score 11; DB 1; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 677 LKINAKLGGIN 687
Db 557 LKINAKLGGIN 567
RESULT 7
ID Q5VXF0 HUMAN PRELIMINARY; PRT; 861 AA.
AC Q5VXF0;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Eukaryotic translation initiation factor 2C, 4.
GN Name=EIP2C4; ORFNames=RP11-239B6.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA White S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Howden P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AL359186; CAH71584.1; -; Genomic DNA.
DR EMBL; AL354864; CAH71584.1; JOINED; Genomic DNA.
DR EMBL; AL354864; CAH73806.1; -; Genomic DNA.
DR EMBL; AL359186; CAH73806.1; JOINED; Genomic DNA.
DR SMR; Q5VXF0; 212-374
DR Ensembl; ENSG00000134698; Homo sapiens.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Pfam; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
KW Initiation factor.
SQ SEQUENCE 861 AA; 97097 MW; F236FF05047534C1 CRC64;
Query Match 1.1%; Score 11; DB 2; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 677 LKINAKLGGIN 687
Db 557 LKINAKLGGIN 567

ID Q4KL66 XENLA PRELIMINARY; PRT; 884 AA.
AC Q4KL66;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BC098982; AAH98982.1; -; mRNA.
DR SMR; Q4KL66; 235-397.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
KW Hypothetical protein.
SQ SEQUENCE 884 AA; 99538 MW; 6416DEB78B4DB2A6 CRC64;
Query Match 1.1%; Score 11; DB 2; Length 884;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 677 LKINAKLGGIN 687
Db 580 LKINAKLGGIN 590

```
RESULT 9
Q4SV66.TETNG
AC Q4SV66.TETNG PRELIMINARY; PRT; 945 AA.
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 21 SCAP13761, whole genome shotgun sequence.
GN ORFNames=GSTENG00012015001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallou O., Aury J.-M., Brunet F., Petit J.-L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Castolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAB01013761; CAF95386.1; -; Genomic_DNA.
DR SNR; Q4SV66; 224-386.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PSS0821; PAZ; 1.
DR PROSITE; PSS0822; PIWI; 1.
SQ SEQUENCE 945 AA; 106438 MW; 9F5E76B38E8D98A1 CRC64;

Query Match 1.1%; Score 11; DB 2; Length 945;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
DB 601 LKINAKLGGIN 611

RESULT 10
Q9NW28 HUMAN
AC Q9NW28
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-MAR-2006, entry version 16.
DE CDNA FLJ10351 fis, clone NT2RM2001141.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Fujimori K.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RN Nat. Genet. 36:40-45(2004).
RN -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK001213; BA991558.1; -; mRNA.
DR Ensembl; ENSG00000197181; Homo sapiens.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PSS0822; PIWI; 1.
SQ SEQUENCE 371 AA; 42215 MW; 59BCF383C997D293 CRC64;

Query Match 1.0%; Score 10; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRQGVSD 805
DB 208 IVVYRQGVSD 217

RESULT 11
Q3TQ88 MOUSE
ID Q3TQ88 MOUSE
AC Q3TQ88
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE 10 days neonate medulla oblongata cDNA, RIKEN full-length enriched
DE library, clone:B830005G15 product:piwi like homolog 2 (Drosophila),
DE full insert sequence. (Fragment).
GN Names=Piwi12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

RA NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX PubMed=16141073; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
Oyama N., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
Banaal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crome M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
Flecher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.P., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
Tamaoka K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Grimmond S.M., Tesdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Pleasy C., Shibata K., Shiraki T., Suzuki S.,
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome."
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT Antisense Transcription in the Mammalian Transcriptome."
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
Sultana R., Takenaka Y., Taylor M.S., Teddale R.D., Tomita M.,
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L.,
Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schriml L.M., Staab F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bono M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara I., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=20530913; PubMed=11076961; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,

RA Muramatsu M., Hayaishizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AKI63647; BAB37436.1; -; mRNA.
DR MGI; MGI:1930036; Pw112.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR003165; Pw1.
DR Pfam; PF02171; Pw1; 1.
DR PROSITE; PSS0822; Pw1; 1.
FT NON TER 1
SQ SEQUENCE 455 AA; 51894 MW; 772A85A184682FCE CRC64;

Query Match 1.0%; Score 10; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
Db |||||||
292 IVVYRDGVSD 301

RESULT 12
Q96SW6 HUMAN PRELIMINARY; PRT; 530 AA.
AC Q96SW6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-MAR-2006, entry version 14.
DE CDNA FLJ14591 fis, clone NT2RM4002034, weakly similar to Homo sapiens
DE hawi mRNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakanatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura M., Shiratori A.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosioki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Inose N.,
RA Musashino K., Yuuki H., Oshino A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Nomiya H., Sato H., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs,"
RL Nat. Genet. 36:40-45 (2004).
CC -----

CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK027497; BAB55155.1; -; mRNA.
DR EMBL; ENSG00000197181; Homo sapiens.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Pw1.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Pw1; 1.
DR PROSITE; PSS0821; PAZ; 1.
DR PROSITE; PSS0822; Pw1; 1.
SQ SEQUENCE 530 AA; 60620 MW; BA4A8AADD809B32F CRC64;

Query Match 1.0%; Score 10; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
Db |||||||
367 IVVYRDGVSD 376

RESULT 13
Q99MV6 MOUSE PRELIMINARY; PRT; 580 AA.
AC Q99MV6;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Pw112.
DE Name=Pw112;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525; DOI=10.1038/86927;
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogonia,"
RL Nat. Genet. 27:422-426 (2001).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF285586; AAK31965.1; -; mRNA.
DR EMBL; ENSMUSG0000003644; Mus musculus.
DR MGI; MGI:1930036; Pw112.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Pw1.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Pw1; 1.
DR PROSITE; PSS0821; PAZ; 1.
DR PROSITE; PSS0822; Pw1; 1.
SQ SEQUENCE 580 AA; 66550 MW; A4EA9D8626FA136 CRC64;

Query Match 1.0%; Score 10; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
Db |||||||
417 IVVYRDGVSD 426

RESULT 14
Q8CDG1 MOUSE PRELIMINARY; PRT; 971 AA.
AC Q8CDG1;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
CC -----

DT 01-MAR-2003, sequence version 1.
 DT 07-FEB-2006, entry version 18.
 DE Adult male testis cDNA, RIKEN full-length enriched library,
 DE clone:4932443D15 product:piwi like homolog 1 (Drosophila)-like, full
 DE insert sequence.
 GN Name=Piwi12;
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kello J., Kitamura H.,
 RA Kitano H., Kollias G., Kriehn S.P., Kruger A., Kumarfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera P., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
 RA Tamoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466681; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gassterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteza G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito K., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Montberns P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [8]

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,
RA Saito R., Saichou H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----
DR EMBL; AK030116; BAC26791.1; -; mRNA.
DR Ensembl; ENSMUSG0000003644; Mus musculus.

DR MGI; MGI:1930036; Pwll2.

DR GO; GO:0005737; C:cytoplasm; IDA.

DR InterPro; IPR003100; PAZ.

DR InterPro; IPR003165; Pwll.

DR Pfam; PF02170; PAZ; 1.

DR Pfam; PF02171; Pwll; 1.

DR PROSITE; PS50821; PAZ; 1.

DR PROSITE; PS50822; Pwll; 1.

SQ SEQUENCE 971 AA; 109458 MW; B63FELL1A2E9DA0F9 CRC64;

Query Match 1.0%; Score 10; DB 2; Length 971;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805

DB 808 IVVYRDGVSD 817

RESULT 15

QJUMB6_MOUSE

ID QJUMB6_MOUSE PRELIMINARY; PRT; 971 AA.

AC QJUMB6;

DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2000, sequence version 1.

DT 07-FEB-2006, entry version 16.

DE MILI.

GN Name=Pwll2; Synonyms=mili;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=21463379; PubMed=11578866; DOI=10.1016/S0925-4773(01)00499-3;

RA Miyagawa S.K., Kimura T., Yomogida K., Kuroiwa A., Tadokoro Y.,

RA Fujita Y., Sato M., Matsuda Y., Nakano T.;

RT "Two mouse pwi-related genes: miwi and mili."

RL Mech. Dev. 108:121-133(2001).

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AB032605; BAA93706.1; -; mRNA.

DR Ensembl; ENSMUSG0000003644; Mus musculus.

DR MGI; MGI:1930036; Pwll2.

DR GO; GO:0005737; C:cytoplasm; IDA.

DR InterPro; IPR003100; PAZ.

DR InterPro; IPR003165; Pwll.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Pwll; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; Pwll; 1.
SQ SEQUENCE 971 AA; 109488 MW; 01E143C6513310FB CRC64;

Query Match 1.0%; Score 10; DB 2; Length 971;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805

DB 808 IVVYRDGVSD 817

Search completed: July 5, 2006, 13:58:50
Job time : 319 secs

THIS PAGE RI ANK 1108701

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:53:42 ; Search time 48 Seconds
(without alignments)
2044.606 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 1020
Sequence: 1 MSSNFPLEKGFYRHSIDP.....RHEMHPLOTNVTKPGMSFA 1020

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

PIR_80:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	45.9	905	2 T23510	hypothetical prote
2	15	1.5	185	2 T27784	hypothetical prote
3	10	1.0	1139	2 T33275	hypothetical prote
4	9	0.9	414	2 A82177	conserved hypothet
5	9	0.9	891	2 T32079	hypothetical prote
6	9	0.9	930	2 A84668	Argonaute (AGO1)-1
7	9	0.9	958	2 S41013	hypothetical prote
8	9	0.9	1000	2 T22391	hypothetical prote
9	9	0.9	1032	2 T23164	hypothetical prote
10	9	0.9	1035	2 T23165	hypothetical prote
11	9	0.9	1040	2 D88568	protein ZK757.3 [i
12	8	0.8	192	2 S59537	heat shock transcr
13	8	0.8	452	1 WZBR4	gene 4 protein - h
14	8	0.8	453	2 B70195	response regulator
15	8	0.8	468	2 A53889	protein-cytosine-p
16	8	0.8	468	2 T43622	targeted effector
17	8	0.8	468	2 S01054	virulence protein
18	8	0.8	477	2 T05202	pectinesterase hom
19	8	0.8	487	2 G95388	probable aldehyde
20	8	0.8	507	2 AE0473	colicin (partial)
21	8	0.8	576	2 S65001	probable membrane
22	8	0.8	988	2 T52134	Zwille protein [im
23	8	0.8	997	2 A84678	Argonaute (AGO1)-1
24	8	0.8	1372	2 F97722	hypothetical prote
25	8	0.8	1372	2 B71724	dna-directed RNA p
26	8	0.8	1526	2 T41522	myosin ii - fisio
27	7	0.7	62	2 C71571	hypothetical prote
28	7	0.7	62	2 E81724	conserved hypothet
29	7	0.7	85	1 A69886	conserved hypothet

30	7	0.7	88	2 G70331	anti sigma factor
31	7	0.7	93	2 A97726	hypothetical prote
32	7	0.7	95	2 T17252	hypothetical prote
33	7	0.7	112	2 E83814	Na/H+ antiporter
34	7	0.7	112	2 S61394	hypothetical prote
35	7	0.7	115	2 H71104	hypothetical prote
36	7	0.7	123	1 HQDVFS	hydrogenase (EC 1.
37	7	0.7	130	2 T08329	hypothetical prote
38	7	0.7	139	1 W6MLB2	E6 protein - bovin
39	7	0.7	150	1 S36991	transposase (clone
40	7	0.7	158	2 W6MLPR	E6 protein - human
41	7	0.7	166	2 F96024	conserved hypothet
42	7	0.7	193	2 B95340	hypothetical prote
43	7	0.7	194	2 B97727	prolyl endopeptida
44	7	0.7	196	2 D84380	imidazoglyceroxol-
45	7	0.7	199	2 C88021	protein M10D9.1 [i

ALIGNMENTS

RESULT 1

T23510
hypothetical protein K08H10.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T23510
R/Gardner, A.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19750
A/Accession: T23510
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-905 <WIL>
A/Cross-References: UNIPROT:Q9XU82; UNIPARC:UPI000017BAA9; EMBL:Z83113; PIDN:CAB05546.1
A/Experimental source: clone K08H10
C/Genetics:
A/Gene: CESP:K08H10.7
A/Map position: 5
A/Intons: 19/3; 86/2; 190/3; 209/2; 269/1; 341/2; 468/3; 671/3; 832/3

Query Match	Best Local Similarity	Score	DB 2;	Length	905;
Matches	468;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	MSSNFPLEKGFYRHSIDPEKMLARPTGKCDGKFKYKTLVLMWPKFSKIYDREYIE	60		
DB	1	MSSNFPLEKGFYRHSIDPEKMLARPTGKCDGKFKYKTLVLMWPKFSKIYDREYIE	60		
QY	61	YEVMTKEVLRNRKGPFPKTEIPIDRAKLFWQHLRHEKKOTDFTLEDVFPDEKTVY	120		
DB	61	YEVMTKEVLRNRKGPFPKTEIPIDRAKLFWQHLRHEKKOTDFTLEDVFPDEKTVY	120		
QY	121	SVCLRLNTVTSKMLVSEKVVKKDSEKDEKLEKILYTMILTYRKKPHLNFSSRENPEKDE	180		
DB	121	SVCLRLNTVTSKMLVSEKVVKKDSEKDEKLEKILYTMILTYRKKPHLNFSSRENPEKDE	180		
QY	181	EANRSTYFLKNVMTQKRYAPFVNEETKVQPAKNFYVDNNSILRVPSFHDNPFESGLE	240		
DB	181	EANRSTYFLKNVMTQKRYAPFVNEETKVQPAKNFYVDNNSILRVPSFHDNPFESGLE	240		
QY	241	VAPRIEMFGIYIGIKELFGEPELVNPAIVDKLFYVAPKMSLDLYLLIYDPOSCHDVR	300		
DB	241	VAPRIEMFGIYIGIKELFGEPELVNPAIVDKLFYVAPKMSLDLYLLIYDPOSCHDVR	300		
QY	301	KDLTKTLMAKMTIRQARPRIRQLLENTLAKCAEVWDMNSRLTERHLTFDLCEBNSL	360		
DB	301	KDLTKTLMAKMTIRQARPRIRQLLENTLAKCAEVWDMNSRLTERHLTFDLCEBNSL	360		
QY	361	VYKVTGSDGRNAKKYDTTLFKIYENKKFIFIEPHPLVLYKVSAGAEIVAPMEHLEVHE	420		
DB	361	VYKVTGSDGRNAKKYDTTLFKIYENKKFIFIEPHPLVLYKVSAGAEIVAPMEHLEVHE	420		

```
Oy      421 KPORYKNRIDLVMDKFLKAATRPDPDYKNTLMLKEIDFSSEELNF  
         |||  
Db      421 KPORYKNRIDLVMDKFLKAATRPDPDYKNTLMLKEIDFSSBELNF
```

RESULT 2

hypochemical protein ZK218.8 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T27784
 R:McMurray, A.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z20418
 A:Accession: T27784
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-185 <WIL>
 A:Cross-references: UNIPROT:O46002; UNIPARC:UPI0000079070; EMBL:Z82085; PIDDN:CAB04988.1
 A:Experimental source: clone ZK218
 C:Genetics:
 A:Gene: CESP:ZK218.8
 A:Map position: 5
 A:Introns: 37/3, 150/3, 174/2

RESULT 3

hypocherical protein R09A1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: J13275
R:Bentley, D., Kemp, K., Harper, M.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid R09A1.
A:Reference number: 21312
A:Accession: J13275
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1139 <BEN>
A:Cross-references: UNIPROT:O61931; UNIPARC:UP1000017B50B; EMBL:AF068711; PIDD:AACT775.
A:Experimental source: strain Bristol N2; clone R09A1
C:Gene: CESP:R09A1.1
A:Map position: 5
A:introns: 314/2; 376/1; 682/2; 774/1; 818/3; 1106/2

RESULT 4

C: conserved hypotheoretical protein VCI624 [imported] - Vibrio cholerae (strain N1691 serogroup O1)

C_Species: Vibrio cholerae

C_Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C_Accession: AB2177

R:Haidelberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Charles, D.; Ermolaeva, V.D.; Vamathavan, U.; Bass, S.; Qin, H.; Dragot, I.; Sellers, F.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A1:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*
A1:Reference number: A82035; MUID:20406833; PMID:10952301

RESULT 5

hypothetical protein T07D3.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T32079
 R:Lamar, B.; Wamsley, P.; Twyman, B.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid T07D3.
 A:Reference number: Z21121
 A:Accession: T32079
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-891 <LAN>
 A:Cross-references: UNIPARC:UPI000002DB7, EMBL:AF016662; PIDD:AA66187.1; GSPDB:GN00022
 A:Experimental source: strain Bristol N2; clone T07D3
 C:Genetics:
 A:Gene: CESP:T07D3..7
 A:Map position: 2
 A:Introns: 34/2, 92/2, 206/1, 560/1, 879/3
 C:Superfamily: rabbit translation initiation factor eIF-2C

RESULT 6

Argonaute (AGO1)-like protein [imported] Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84668
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
neus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10517197
A:Accession: A84668
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-930 <STO>
A:Cross-references: UNIPROT:Q9ZVD5; UNIPARC:UPI0000179101; GB:AE002093; F
C:Gene: At2g27040

A;Map position: 2
C;Superfamily: rabbit translation initiation factor eIF-2C

Query Match 0.9%; Score 9; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGG 685
|||||
Db 631 LKINAKLGG 639

RESULT 7
S41013
hypothetical protein ZK757.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Dec-2004
C;Accession: S41013
R;Thomas, K.
submitted to the EMBL Data Library, December 1993
A;Reference number: S41011
A;Accession: S41013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-958 <THO>
A;Cross-references: UNIPROT:Q9TW94; UNIPARC:UPI0000179102; EMBL:Z29121
C;Genetics:
A;Introns: 52/2; 350/2; 422/3; 592/1; 679/3; 721/2; 887/2; 929/1

Query Match 0.9%; Score 9; DB 2; Length 958;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 715 VGIDVTHPT 723
|||||
Db 658 VGIDVTHPT 666

RESULT 8
T22391
hypothetical protein F48F7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22391; T23234
R;Coles, L.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19559
A;Accession: T22391
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1000 <WIL>
A;Cross-references: UNIPROT:Q20578; UNIPARC:UPI0000179103; EMBL:Z69661; PIDN:CAA93496.1;
A;Experimental source: clone F48F7
R;Coles, L.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19712
A;Accession: T23234
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1000 <W12>
A;Cross-references: UNIPARC:UPI0000179103; EMBL:Z69663; PIDN:CAA93512.1; GSPDB:GN000028;
A;Experimental source: clone K02B9
C;Genetics:
A;Gene: CESP:F48F7.1
A;Map position: X
A;Introns: 70/3; 128/2; 185/2; 673/1
C;Superfamily: rabbit translation initiation factor eIF-2C

Query Match 0.9%; Score 9; DB 2; Length 1000;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IIVYRDGVS 804

Db 807 IIVYRDGVS 815
|||||

RESULT 9
T23164
hypothetical protein T22B3.2a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23164; T25099
R;Cottage, A.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19701
A;Accession: T23164
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1032 <WIL>
A;Cross-references: UNIPROT:Q21079; UNIPARC:UPI000002A1C2; EMBL:Z68750; PIDN:CAA92969.1
A;Experimental source: clone K01A6
R;Lennard, N.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19981
A;Accession: T25099
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1032 <W12>
A;Cross-references: UNIPARC:UPI000002A1C2; EMBL:Z68300; PIDN:CAA92618.1; GSPDB:GN000022;
A;Experimental source: clone T22B3
C;Genetics:
A;Gene: CESP:T22B3.2a
A;Map position: 4
A;Introns: 19/1; 71/3; 129/2; 424/2; 496/3; 607/1; 666/1; 753/3; 795/2; 961/2; 1003/1
C;Superfamily: rabbit translation initiation factor eIF-2C

Query Match 0.9%; Score 9; DB 2; Length 1032;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 715 VGIDVTHPT 723
|||||
Db 732 VGIDVTHPT 740

RESULT 10
T23165
hypothetical protein T22B3.2b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23165; T25100
R;Cottage, A.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19701
A;Accession: T23165
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1035 <WIL>
A;Cross-references: UNIPROT:Q21079; UNIPARC:UPI000008666E; EMBL:Z68750; PIDN:CAA92970.1
A;Experimental source: clone K01A6
R;Lennard, N.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19981
A;Accession: T25100
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1035 <W12>
A;Cross-references: UNIPARC:UPI000008666E; EMBL:Z68300; PIDN:CAA92619.1; GSPDB:GN000022;
A;Experimental source: clone T22B3
C;Genetics:
A;Gene: CESP:T22B3.2b
A;Map position: 4
A;Introns: 19/1; 74/3; 132/2; 427/2; 499/3; 610/1; 669/1; 756/3; 798/2; 964/2; 1006/1
C;Superfamily: rabbit translation initiation factor eIF-2C

Query Match 0.9%; Score 9; DB 2; Length 1035;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0

QY 715 VGIDVTHPT 723
|||||
Db 735 VGIDVTHPT 743

RESULT 11
D88568
protein ZK757.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D88568
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: D88568
A:Status: preliminary
A:Keywords: preliminary
A:Molecule type: DNA
A:Residues: 1-1040 <STO>
A:Cross-references: UNIPROT:P34681; UNIPARC:UPI000013BC28; GB:chr_III; PIDN:CAA82941.1;
C:Genetics:
A:Gene: ZK757.3
A:Map position: 3
C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 0.9%; Score 9; DB 2; Length 1040;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0

QY 715 VGIDVTHPT 723
|||||
Db 740 VGIDVTHPT 748

RESULT 12
S59537
heat shock transcription factor HSP21 - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S59537; S52303
R:Czarnecka-Verner, E.; Yuan, C.X.; Fox, P.C.; Gurley, W.B.
Plant Mol. Biol. 29, 37-51, 1995
A:Title: Isolation and characterization of six heat shock transcription factor cDNA clones
A:Reference number: S59537; MUID:96017612; PMID:7579166
A:Accession: S59537
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-192 <CZA>
A:Cross-references: UNIPROT:Q43454; UNIPARC:UPI0000177D9B; EMBL:Z46952
C:Genetics:
A:Gene: HSP21
C:Superfamily: tomato heat shock transcription factor HSP24; HSF DNA-binding domain homolog
C:Keywords: DNA binding; heat shock; leucine zipper; nucleus; stress-induced protein; transcription factor
F:46-141/Domain: HSF DNA-binding domain homology <HSP>

Query Match 0.8%; Score 8; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.5; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 627 RQLNAYGF 634
|||||
Db 99 RQLNAYGF 106

RESULT 13
WZBE4
gene 4 protein - human herpesvirus 3

Query Match 0.8%; Score 8; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 18; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 325 LLENLKLK 332
|||||
Db 321 LLENLKLK 328

RESULT 14
B70195
response regulatory protein (rrp-2) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 02-Jun-2003
C:Accession: B70195
R:Praser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whiston, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugl; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horat, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70195
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-453 <KLE>
A:Cross-references: UNIPARC:UPI00000575DE; GB:AE001176; GB:AE000783; NID:g2688699; PIDN:G2688699
A:Experimental source: strain B31
C:Superfamily: response regulator, NtrC type; response regulator homology; RNA polymerase
C:Keywords: phosphoprotein
F:6-115/Domain: response regulator homology <RRH>
F:147-369/Domain: RNA polymerase sigma factor interaction domain homology <SFI>
F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 0.8%; Score 8; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 18; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 694 EIAEISPE 701
|||||
Db 247 EIAEISPE 254

RESULT 15
A53889
protein-tyrosine-phosphatase (EC 3.1.3.48) Yop51 - Yersinia enterocolitica (strain W227)
N:Alternate names: virulence protein Yop51
C:Species: Yersinia enterocolitica
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: A53889
R:Michiels, T.; Cornelis, G.
Microb. Pathog. 5, 449-459, 1988
A:Title: Nucleotide sequence and transcription analysis of yop51 from Yersinia enterocolitica
A:Reference number: A53889; MUID:89218619; PMID:3244311
A:Accession: A53889
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-468 <MIC>
A;Cross-references: UNIPROT:P15273; UNIPARC:UPI00000000256; GB:M30457; NID:gl55530; PIDN:
C;Genetics:
A;Gene: Yop51
A;Genome: plasmid
C;Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase
Query Match 0.8%; Score 8; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 655 VTKALASL 662
| | | | | | | |
Db 364 VTKALASL 371
Search completed: July 5, 2006, 13:59:43
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM protein - protein search, using sw model

Run on: July 5, 2006, 13:49:57 ; Search time 196 Seconds
(without alignments)
2379.392 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 1020
Sequence: 1 MSNPFLEKGFVHSLDPE.....RHEMHZLQTNVVKPGMSFA 1020

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2589342

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq.8.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	1020	4	Aau01856 C. elegans
2	11	1.1	861	10	Aef73792 Human Arg
3	11	1.1	924	7	Adj69189 Human hea
4	11	1.1	924	8	Adj94699 Human euk
5	10	1.0	371	4	Aab27222 Human pro
6	10	1.0	530	4	Aab94209 Human pro
7	10	1.0	530	7	Adj69986 Human hea
8	10	1.0	580	4	Aau07866 Polypepti
9	10	1.0	678	8	Adx66443 Plant ful
10	9	0.9	94	10	Aef73781 Caenorhab
11	9	0.9	94	10	Aef73783 Caenorhab
12	9	0.9	94	10	Aef73782 Caenorhab
13	9	0.9	389	6	Adb17506 Wheat pos
14	9	0.9	389	9	Aec75730 Wheat Arg
15	9	0.9	850	8	Adq80691 Arabidops
16	9	0.9	892	6	Adq80689 Arabidops
17	9	0.9	901	6	Adb17464 Corn post
18	9	0.9	901	9	Aec75688 Corn post
19	9	0.9	904	6	Adb17511 Rice post
20	9	0.9	904	9	Aec75735 Rice post
21	9	0.9	915	6	Adb17496 Soybean p
22	9	0.9	915	9	Aec75720 Soybean p
23	9	0.9	916	6	Adb17498 Wheat pos

24	9	0.9	916	9	Aec75722 Wheat Arg
25	9	0.9	924	8	Adq80687 Arabidops
26	9	0.9	930	7	Adq31435 Plant yie
27	9	0.9	930	8	Adi44099 Plant tra
28	8	0.8	47	3	Aag50116 Arabidops
29	8	0.8	108	10	Aef73786 Arabidops
30	8	0.8	108	10	Aef73787 Arabidops
31	8	0.8	108	10	Aef73784 Rice Argo
32	8	0.8	108	10	Aef73785 Rice Argo
33	8	0.8	109	4	Aau53734 Propionib
34	8	0.8	109	6	Abm50253 Propionib
35	8	0.8	121	6	Adm06608 Allicoec
36	8	0.8	126	7	Adk52503 Penicilli
37	8	0.8	161	8	Adri15803 Kinase ly
38	8	0.8	183	10	Aef14816 RNA-depen
39	8	0.8	235	8	Adr28077 NPB polyp
40	8	0.8	243	4	Aab59398 Yarsinia
41	8	0.8	248	8	Ady07084 Plant ful
42	8	0.8	287	4	Aag78296 Yarsinia
43	8	0.8	306	10	Aef14819 RNA-depen
44	8	0.8	326	3	Aag52610 Arabidops
45	8	0.8	337	8	Adx77635 Plant ful

ALIGNMENTS

RESULT 1
AAU01856
ID AAU01856 standard; protein; 1020 AA.
XX
AC AAU01856;
XX
DT 07-SEP-2001 (first entry)
XX
DE C. elegans RNA interference pathway protein RDE-1.
XX
KW RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;
KW double-stranded RNA-dependent gene silencing.
XX
OS Caenorhabditis elegans.
XX
PN WO200129058-A1.
XX
PD 26-APR-2001.
XX
PF 13-OCT-2000; 2000WO-US028470.
XX
PR 15-OCT-1999; 99US-0159776P.
PR 30-MAR-2000; 2000US-0193218P.
(UYMA-) UNIV MASSACHUSETTS.
(CARN-) CARNEGIE INST WASHINGTON.
Mello CC, Fire A, Tabara H, Grishok A;
WPI; 2001-316239/33.
N-PSDB; AAS03282, AAS03283.
Novel RNA interference pathway genes and their protein products involved in mediation of genetic interference, useful for modulating and studying regulation of RNA interference pathway.
Claim 3; Fig 6; 76pp; English.

The sequence represents the RNA interference (RNAi) pathway protein RDE-1. RDE-1 and RDE-4 are involved in the pathway mediating double-stranded RNA-dependent gene silencing (genetic interference) RDE-1 and RDE-4 protein is useful for preparing an RNAi agent, by incubating a dsRNA in the presence of the proteins. The prepared RNAi agents can be used as sequence-specific interfering agents for targeted genetic interference. The nucleic acids are useful for studying the regulation of RNAi pathway and to generate knockout strains of animals such as C.elegans. RDE-1 and

	CC	RDE-4 genes and their products are useful for modulating RNAi pathway activity. The polypeptides are useful for generating and testing antibodies specific for the polypeptides which are useful for studying the RNAi pathway in C.elegans and other organisms. RNAi pathway genes are useful for mediating specific processes, e.g. a gene that mediates dsRNA uptake by the cells is useful for transporting other RNAs into cells or for facilitating entry of agents such as drugs into cells.
	CC	RNAi pathway mutant strains (rde-1, rde-4 mutants) are useful in genetic screens to identify additional RNAi pathway components
	XX	
	SQ	Sequence 1020 AA;
		Query Match 100.0%; Score 1020; DB 4; Length 1020; Best Local Similarity 100.0%; Pred. No. 0; Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1	MSSNFPPELEKGFYRHSLDPEMKWLARPTGCKDGKEYKKVLLLVNMFKFSSKIYDREYYE 60
Dd	1	MSSNFPPELEKGFYRHSLDPEMKWLARPTGCKDGKEYKKVLLLVNMFKFSSKIYDREYYE 60
Qy	61	YEVMKTKEVLNRKPGRPKPTEIPIPRAKLFWOHLRHEKKQTDFILEDVVFDSDKTVY 120
Dd	61	YEVMKTKEVLNRKPGRPKPTEIPIPRAKLFWOHLRHEKKQTDFILEDVVFDSDKTVY 120
Qy	121	SVCRRLNTVTSKMLYSEKVVKVDSEKKDKOLEKKILYTMLTYRRKFHNFSENPEKDE 180
Dd	121	SVCRRLNTVTSKMLYSEKVVKVDSEKKDKOLEKKILYTMLTYRRKFHNFSENPEKDE 180
Qy	181	EANRSYKFLKNVTQKYRYAPFNBEIIKVFPAKNFYDNNSILRVPSFHDPNRFQSLE 240
Dd	181	EANRSYKFLKNVTQKYRYAPFNBEIIKVFPAKNFYDNNSILRVPSFHDPNRFQSLE 240
Qy	241	VAPIRIEAWFGIYGIKELPDGEPLNFAIVDKLFYNAPKMSLLDYLLIIVDQSCNDVVR 300
Dd	241	VAPIRIEAWFGIYGIKELPDGEPLNFAIVDKLFYNAPKMSLLDYLLIIVDQSCNDVVR 300
Qy	301	KDLTKTLMAGKWITIRQAARPRIQLLENLKLCAEVWDNEMSRLTERHLTFLDLCENSL 360
Dd	301	KDLTKTLMAGKWITIRQAARPRIQLLENLKLCAEVWDNEMSRLTERHLTFLDLCENSL 360
Qy	361	VYKVGSGDRGNNAKGYDTTLFKYIEENKKETIEPHLPFLVKVKSAAKAYAVPMEHLEVHE 420
Dd	361	VYKVGSGDRGNNAKGYDTTLFKYIEENKKETIEPHLPFLVKVKSAAKAYAVPMEHLEVHE 420
Qy	421	KPORYKNIIDIAVMDDKFLKRATRAPHDYKENTLKMCLKEDSPSSEELNVFERPGLSKLQM 480
Dd	421	KPORYKNIIDIAVMDDKFLKRATRAPHDYKENTLKMCLKEDSPSSEELNVFERPGLSKLQM 480
Qy	481	IECPGKVLKEPMLVNVSNEIQIMTPVIRGFOSKQLNVVPVEKELCCAFFVNETAGNPCLE 540
Dd	481	IECPGKVLKEPMLVNVSNEIQIMTPVIRGFOSKQLNVVPVEKELCCAFFVNETAGNPCLE 540
Qy	541	ENDVVKFTTELIGGCKFRGIIRIGANENRGASIMYDATNKNEYAFYQNCTLTNGIGRFEIA 600
Dd	541	ENDVVKFTTELIGGCKFRGIIRIGANENRGASIMYDATNKNEYAFYQNCTLTNGIGRFEIA 600
Qy	601	ATEAKNNPERLPDKEQKVLMTFIISIKROLNAGFYVKHYCDHTTIGVANOHITSETVTKALA 660
Dd	601	ATEAKNNPERLPDKEQKVLMTFIISIKROLNAGFYVKHYCDHTTIGVANOHITSETVTKALA 660
Qy	661	SLRHEKGSKRIFYQIALKINAKLGGINOLDWSEIAEISPEEKERRKTMPLTWYVGIDVT 720
Dd	661	SLRHEKGSKRIFYQIALKINAKLGGINOLDWSEIAEISPEEKERRKTMPLTWYVGIDVT 720
Qy	721	HPTSYSGIDYSIAA VVASINPGGTIYRNMI VTQEBCRPGERAVA HGRERTDILEAKFVKL 780
Dd	721	HPTSYSGIDYSIAA VVASINPGGTIYRNMI VTQEBCRPGERAVA HGRERTDILEAKFVKL 780
Qy	781	LRFEAENNDRAPAHI VVYRDGVSDSEM LRVSHDELRSI KSEVKQFMFSERGEDPEPKYT 840
Dd	781	LRFEAENNDRAPAHI VVYRDGVSDSEM LRVSHDELRSI KSEVKQFMFSERGEDPEPKYT 840
Qy	841	FVIQKRHNTLRMRMEXDKPVWNKDLP TAETDVAAVAVKQWESDMKSKETGIVNPSSG 900

three-dimensional molecular modeling algorithm to the atomic coordinates of an Argonaute protein to determine the spatial coordinates of the binding pocket of the Argonaute protein, and electronically screening the stored spatial coordinates of a set of candidate agents against the spatial coordinates of the Argonaute protein binding pocket; a computer-based method for the analysis of the interaction of a molecular structure with an Argonaute protein; a data array comprising the atomic coordinates of an Argonaute protein; a computer-readable storage medium encoded with the atomic coordinates of an Argonaute protein; an electronic representation of a domain or binding site of the Argonaute protein; a method for obtaining a crystal of an Argonaute protein; a method for identifying an agent that modulates the activity of an RNAi construct by identifying an agent that modulates the expression and/or activity of an Argonaute protein; a method for identifying an agent that potentiates the activity of an RNAi construct by identifying an agent that increases the expression and/or activity of an Argonaute protein; a method of identifying an agent that modulates the activity of an RNAi construct by examining a change in Argonaute protein activity in the presence of a candidate agent; a composition for targeted gene inhibition comprising an agent that modulates the RNase activity of an Argonaute protein; a cell line that overexpresses an Argonaute protein; an assay for identifying nucleic acid sequences for conferring a particular phenotype in a cell; a nucleic acid composition composed of a nucleic acid comprising an RNAi construct and a nucleic acid encoding an Argonaute protein, where the RNAi construct comprises a nucleic sequence encoding a single-strand short interfering RNA (siRNA); a pharmaceutical composition comprising the nucleic acid composition; and a cell expressing the nucleic acid composition. The methods and compositions of the invention are useful for enhancing the effectiveness of an RNAi therapeutic.

XX SQ Sequence 861 AA;

Query Match 1.1%; Score 11; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
|||||
Db 557 LKINAKLGGIN 567

RESULT 3

ADJ69189
ID ADJ69189 standard; protein; 924 AA.

AC ADJ69189;

XX 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SeqID995.

XX mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.

XX Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003WO-US010870.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-0389987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX WPI; 2003-845369/78.
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.

XX Claim 1; SEQ ID NO 995; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nontropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX SQ Sequence 924 AA;

Query Match 1.1%; Score 11; DB 7; Length 924;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
|||||
Db 620 LKINAKLGGIN 630

RESULT 4

ADJ94699

ID ADJ94699 standard; protein; 924 AA.

AC ADJ94699;

XX 06-MAY-2004 (first entry)

XX Human eukaryotic translation initiation factor 2C3 (eIF2C3) protein.

XX cytostatic; antiinflammatory; virucide; immunosuppressive; tumour;
KW inflammatory; infectious disease; viral infection; degenerative;
KW autoimmune; gene therapy; Argonaute family;
KW eukaryotic translation initiation factor 2C3; eIF2C3; human.

XX Homo sapiens.

XX WO2004007718-A2.

XX 22-JAN-2004.

XX 10-JUL-2003; 2003WO-EP007516.

XX 10-JUL-2002; 2002EP-00015532.

XX 23-AUG-2002; 2002EP-00018906.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Tuschl T, Martinez J, Patkaniowska A, Urlaub H, Luehrmann R;

XX WPI; 2004-122948/12.

XX N-PSDB; ADJ94705.

PT New single-stranded RNA molecule having a length from 14-50 nucleotides,
 PT useful for preventing or treating tumor, inflammatory, infectious, e.g.
 PT viral infections, degenerative and autoimmune diseases.

XX Example; Fig 14; 73pp; English.

XX The invention relates to a novel single-stranded RNA molecule having a
 CC length from 14-50 nucleotides where at least 14-20 of the 5'-most
 CC nucleotides are substantially complementary to a target transcript. The
 CC RNA molecule of the invention demonstrates cytostatic, antiinflammatory,
 CC virucide and immunosuppressive activities and may be useful for
 CC inhibiting the expression of a target gene in vitro or in vivo,
 CC preferably for preventing or treating diseases associated with the
 CC overexpression of at least one target transcript. The diseases may be
 CC selected from tumour diseases, inflammatory diseases, infectious diseases
 CC such as viral infections, degenerative diseases and autoimmune diseases.
 CC Furthermore, the molecules of the invention may be utilised during gene
 CC therapy. The current sequence is that of the human eukaryotic translation
 CC initiation factor 2C3 (eIF2C3) protein of the invention.

XX Sequence 924 AA;

Query Match 1.1%; Score 11; DB 8; Length 924;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGGIN 687
 |||||
 Db 620 LKINAKLGGIN 630

RESULT 5
 AAB92722
 ID AAB92722 standard; protein; 371 AA.

AC AAB92722;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:11144.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 11144; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX Sequence 371 AA;

Query Match 1.0%; Score 10; DB 4; Length 371;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
 |||||

Db 208 IVVYRDGVSD 217

RESULT 6
 AAB94209
 ID AAB94209 standard; protein; 530 AA.

XX AAB94209;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:14559.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 14559; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 530 AA;

Query Match 1.0%; Score 10; DB 4; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
 |||||
 Db 367 IVVYRDGVSD 376

RESULT 7
 ADJ69986
 ID ADJ69986 standard; protein; 530 AA.
 AC
 AC ADJ69986;
 XX
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID1792.
 XX
 XX mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 XX 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 XX (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 PA
 XX
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 PI
 XX WPI; 2003-845369/78.
 DR
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,

PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX

PS Claim 1; SEQ ID NO 1792; 180pp; English.

XX
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, neurotropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX

SQ Sequence 530 AA;

Query Match 1.0%; Score 10; DB 7; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
 |||||
 Db 367 IVVYRDGVSD 376

RESULT 8
 AAU07866
 ID AAU07866 standard; protein; 580 AA.
 XX
 AC AAU07866;
 XX
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Polypeptide sequence for mammalian Spg16.
 XX
 XX Mammalian; reproductive-specific protein; male infertility;
 KW spermatogenesis; sperm count disorder; anti infertility; reproduction.
 KW Mammalia.
 OS
 XX WO200166752-A2.
 XX
 XX
 PD 13-SEP-2001.
 XX
 XX
 PF 07-MAR-2001; 2001WO-US007371.
 XX
 XX
 PR 07-MAR-2000; 2000US-0187518P.
 PR 12-JAN-2001; 2001US-0261557P.
 XX
 XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA
 XX
 XX Wang PJ, Page DC;
 XX
 XX WPI; 2001-570774/64.
 DR
 DR N-PSDB; AAS13630.
 XX
 XX Novel reproduction-specific protein, useful for treating disorders of
 PT reduced sperm count, enhancing/increasing sperm count and/or sperm
 PT activity.
 XX
 XX Claim 22; Fig 16; 151pp; English.
 XX
 CC The present invention relates to the isolation of novel mammalian and
 CC human reproductive-specific proteins (AAU07859-AAU07899), and the nucleic
 CC acids encoding them. The nucleic acids encoding reproductive-specific
 CC proteins are useful for diagnosing infertility which is a result of

CC reduced sperm count, reduced sperm motility, malformed sperm or
 CC combinations of these. The sequences of the invention are useful as
 CC markers for spermatogonial cells, for identifying genes or proteins
 CC characteristic of male infertility, diagnosing or aiding in the diagnosis
 CC of infertility in men, and for contraception in which sperm production or
 CC sperm count is reduced or defective sperm is produced. Antibodies to
 CC reproductive-specific proteins are useful for determining the presence of
 CC these proteins in a sample obtained from a man being assessed for
 CC infertility, for identifying the expression of genes in particular cell
 CC type or particular developmental stage, for studies of spermatogenesis,
 CC and for immunofluorescence of germ cells or in Western blots for
 CC assessing the presence of the protein the antibody binds. The sequences
 CC of the invention are also useful for treating disorders of reduced sperm
 CC count, and for increasing sperm count and/or sperm activity. The nucleic
 CC acids of the invention are useful in gene therapy. AAU07859-AAU07882
 CC represent the mammalian reproduction-specific proteins of the present
 CC invention
 XX
 SQ Sequence 580 AA;

Query Match 1.0%; Score 10; DB 4; Length 580;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRDGVSD 805
 DB 417 IVVYRDGVSD 426
 |||||

RESULT 9
 ADX66443
 ID ADX66443 standard; protein; 678 AA.
 XX
 AC ADX66443;

21-APR-2005 (first entry)

Plant full length insert polypeptide seqid 37286.

plant protectant; plant growth regulant; gene therapy; plant;
 recombinant DNA construct; physical array; plant breeding marker;
 cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 extreme osmotic condition; pathogen tolerance; pest tolerance;
 growth rate; cell cycle pathway; disease resistance;
 Galactanmannan production; lignin production; plant growth regulator;
 yield; plant growth; plant development; seed oil; protein yield;
 protein content.

Unidentified.

US2004034888-A1.

19-FEB-2004.

28-APR-2003; 2003US-00425114.

06-MAY-1999; 99US-00304517.

05-NOV-2001; 2001US-00985678.

(LIU//) LIU J.
 (ZHOU//) ZHOU Y.
 (KOVA//) KOVALIC D K.
 (SCRE//) SCREEN S E.
 (TAB//) TABASKA J E.
 (CAOY//) CAO Y.

Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 WPI; 2004-180133/17.

New recombinant DNA construct, useful for improving plant tolerance to
 cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 pests, for conferring increased resistance to plant disease, or for

PT improving yield.
 XX Claim 1; SEQ ID NO 37286; 15pp; English.
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactanmannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 678 AA;

Query Match 1.0%; Score 10; DB 8; Length 678;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 731 SIAAVVASIN 740
 DB 440 SIAAVVASIN 449
 |||||

RESULT 10

AEF73781

ID AEF73781 standard; protein; 94 AA.

XX AEF73781;

06-APR-2006 (first entry)

Caenorhabditis briggsae Argonaute protein Piwi domain polypeptide.

Argonaute; RNA interference; RNAi; drug screening; X-ray crystallography;
 gene silencing.

Caenorhabditis briggsae.

WO2006015258-A2.

09-FEB-2006.

28-JUL-2005; 2005WO-US027084.

28-JUL-2004; 2004US-0592297P.

29-JUL-2004; 2004US-0592269P.

(COLD-) COLD SPRING HARBOR LAB.

Joshua-Tor L, Song J, Hannon GJ, Liu J, Carmell MA, Rivas F;
 Marsden C;

WPI; 2006-155768/16.

Crystalline Argonaute useful for identifying agent that binds Argonaute
 protein, utilized for treating pancreatic cancer, treating hepatitis C
 infection, cancer or inflammatory diseases.

Example 10; Fig 21; 215pp; English.

The present invention relates to Argonaute proteins, which are involved
 in RNA interference (RNAi). The invention provides a crystalline

CC Argonaute protein, which comprises (i) an N-terminus, middle and PIWI
 CC domain which form a crescent-shaped base, and (ii) a PAZ domain, which is
 CC positioned above the crescent-shaped base, resulting in a cleft between
 CC the crescent-shaped base and the PAZ domain. The structure of the full-
 CC length Argonaute protein AEF73751 from the archaeobacterium Pyrococcus
 CC furiosus (PfAgol), as determined by x-ray crystallography to 2.25 Angstrom
 CC resolution, is provided. The invention also provides an isolated complex
 CC comprising an Argonaute protein and a single-stranded RNA hybridized to
 CC its target nucleic acid. In certain embodiments, the single-stranded RNA
 CC is bound to the PAZ domain of the Argonaute protein, and may further
 CC interact with the crescent-shaped base of the Argonaute protein. Also
 CC claimed are: a method of determining the three-dimensional structure of
 CC an Argonaute protein or its mutant, derivative, variant, analog, homolog,
 CC sub-domain or fragment by alignment with the PfAgol amino acid sequence to
 CC match homologous regions; a method of identifying an agent that modulates the
 CC Argonaute protein by applying a three-dimensional molecular modeling
 CC algorithm to the atomic coordinates of an Argonaute protein to determine
 CC the spatial coordinates of the binding pocket of the Argonaute protein,
 CC and electronically screening the stored spatial coordinates of a set of
 CC candidate agents against the spatial coordinates of the Argonaute protein
 CC binding pocket; a computer-based method for the analysis of the
 CC interaction of a molecular structure with an Argonaute protein; a data
 CC array comprising the atomic coordinates of an Argonaute protein; a
 CC computer-readable storage medium encoded with the atomic coordinates of
 CC an Argonaute protein; an electronic representation of a domain or binding
 CC site of the Argonaute protein; a method for obtaining a crystal of an
 CC Argonaute protein; a method for identifying an agent that modulates the
 CC activity of an RNAi construct by identifying an agent that modulates the
 CC expression and/or activity of an Argonaute protein; a method for
 CC identifying an agent that potentiates the activity of an RNAi construct
 CC by identifying an agent that increases the expression and/or activity of
 CC an Argonaute protein; a method of identifying an agent that modulates the
 CC activity of an RNAi construct by examining a change in Argonaute protein
 CC activity in the presence of a candidate agent; a composition for targeted
 CC gene inhibition comprising an agent that modulates the RNase activity of
 CC an Argonaute protein; a cell line that overexpresses an Argonaute protein
 CC ; an assay for identifying nucleic acid sequences for conferring a
 CC particular phenotype in a cell; a nucleic acid composition composed of a
 CC nucleic acid comprising an RNAi construct and a nucleic acid encoding an
 CC Argonaute protein, where the RNAi construct comprises a nucleic sequence
 CC encoding a single-strand short interfering RNA (siRNA); a pharmaceutical
 CC composition comprising the nucleic acid composition; and a cell
 CC expressing the nucleic acid composition. The methods and compositions of
 CC the invention are useful for enhancing the effectiveness of an RNAi
 CC therapeutic. The present sequence is that of a *Caenorhabditis briggsae*
 CC Argonaute protein PIWI domain polypeptide, which was included in an
 CC alignment of Ago protein polypeptides AEF73774-AEF73788 to examine
 CC conservation of active site residues.

XX Sequence 94 AA;

Query Match 0.9%; Score 9; DB 10; Length 94;

Best Local Similarity 100.0%; Pred. No. 3.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 IVVYRQGV 804

Db 70 IVVYRQGV 78

|||||||

RESULT 11

AEF73783

ID AEF73783 standard; protein; 94 AA.

XX AEF73783;

AC AEF73783;

XX 06-APR-2006 (first entry)

DT 06-APR-2006 (first entry)

XX *Caenorhabditis elegans* ALG2 Argonaute protein PIWI domain polypeptide.DE *Caenorhabditis elegans* ALG2 Argonaute protein PIWI domain polypeptide.

XX Argonaute; RNA interference; RNAi; drug screening; X-ray crystallography;

KW gene silencing.

XX

OS *Caenorhabditis briggsae*.

XX WO2006015258-A2.

XX 09-FEB-2006.

XX 28-JUL-2005; 2005WO-US027084.

XX 28-JUL-2004; 2004US-0592297P.

XX 29-JUL-2004; 2004US-0592269P.

XX (COLD-) COLD SPRING HARBOR LAB.

XX Joshua-Tor L, Song J, Hannon GJ, Liu J, Carmell MA, Rivas F;

XX Marsden C;

XX WPI; 2006-155768/16.

XX Crystalline Argonaute useful for identifying agent that binds Argonaute

XX protein, utilized for treating pancreatic cancer, treating hepatitis C

XX infection, cancer or inflammatory diseases.

XX Example 10; Fig 21; 215pp; English.

XX The present invention relates to Argonaute proteins, which are involved

XX in RNA interference (RNAi). The invention provides a crystalline

XX Argonaute protein, which comprises (i) an N-terminus, middle and PIWI

XX domain which form a crescent-shaped base, and (ii) a PAZ domain, which is

XX positioned above the crescent-shaped base, resulting in a cleft between

XX the crescent-shaped base and the PAZ domain. The structure of the full-

XX length Argonaute protein AEF73751 from the archaeobacterium *Pyrococcus*XX *furiosus* (PfAgol), as determined by x-ray crystallography to 2.25 Angstrom

XX resolution, is provided. The invention also provides an isolated complex

XX comprising an Argonaute protein and a single-stranded RNA hybridized to

XX its target nucleic acid. In certain embodiments, the single-stranded RNA

XX is bound to the PAZ domain of the Argonaute protein, and may further

XX interact with the crescent-shaped base of the Argonaute protein. Also

XX claimed are: a method of determining the three-dimensional structure of

XX an Argonaute protein or its mutant, derivative, variant, analog, homolog,

XX sub-domain or fragment by alignment with the PfAgol amino acid sequence to

XX match homologous regions; a method of identifying an agent that binds an

XX Argonaute protein by applying a three-dimensional molecular modeling

XX algorithm to the atomic coordinates of an Argonaute protein to determine

XX the spatial coordinates of the binding pocket of the Argonaute protein,

XX and electronically screening the stored spatial coordinates of a set of

XX candidate agents against the spatial coordinates of the Argonaute protein

XX binding pocket; a computer-based method for the analysis of the

XX interaction of a molecular structure with an Argonaute protein; a data

XX array comprising the atomic coordinates of an Argonaute protein; a

XX computer-readable storage medium encoded with the atomic coordinates of

XX an Argonaute protein; an electronic representation of a domain or binding

XX site of the Argonaute protein; a method for obtaining a crystal of an

XX Argonaute protein; a method for identifying an agent that modulates the

XX activity of an RNAi construct by identifying an agent that modulates the

XX expression and/or activity of an Argonaute protein; a method for

XX identifying an agent that potentiates the activity of an RNAi construct

XX by identifying an agent that increases the expression and/or activity of

XX an Argonaute protein; a method of identifying an agent that modulates the

XX activity of an RNAi construct by examining a change in Argonaute protein

XX activity in the presence of a candidate agent; a composition for targeted

XX gene inhibition comprising an agent that modulates the RNase activity of

XX an Argonaute protein; a cell line that overexpresses an Argonaute protein

XX ; an assay for identifying nucleic acid sequences for conferring a

XX particular phenotype in a cell; a nucleic acid composition composed of a

XX nucleic acid comprising an RNAi construct and a nucleic acid encoding an

XX Argonaute protein, where the RNAi construct comprises a nucleic sequence

XX encoding a single-strand short interfering RNA (siRNA); a pharmaceutical

XX composition comprising the nucleic acid composition; and a cell

XX expressing the nucleic acid composition. The methods and compositions of

XX the invention are useful for enhancing the effectiveness of an RNAi

XX therapeutic. The present sequence is that of a *Caenorhabditis elegans*

XX ALG2 Argonaute protein PIWI domain polypeptide, which was included in an

XX alignment of Ago protein polypeptides AEF73774-AEF73788 to examine

XX conservation of active site residues.

DR N-PSDB; ADB17505.
 XX
 PT New polynucleotides encoding polypeptides comprising post-transcriptional
 PT gene silencing activity useful in creating transgenic plants having
 PT enhanced or deficient post-transcriptional gene silencing.
 XX
 PS
 XX Disclosure; Page 85-86; 107pp; English.
 XX
 CC The invention discloses an isolated polynucleotide encoding a polypeptide
 CC having post-transcriptional gene silencing (PTGS) activity. PTGS operates
 CC at the level of sequence specific RNA degradation and down regulates
 CC transgene expression in plants. Also claimed is a recombinant DNA
 CC construct comprising the polynucleotide operably linked to at least one
 CC regulatory sequence, transforming a cell with the polynucleotide, a cell
 CC comprising the recombinant construct, producing a polypeptide having post
 CC -transcriptional gene silencing activity, producing a plant, a plant and
 CC polypeptide comprising post-transcriptional gene silencing activity. Also
 CC disclosed are methods for obtaining a nucleic acid fragment encoding all,
 CC or a substantial portion of, the amino acid sequence encoding a
 CC polypeptide involved in post-transcriptional gene silencing, positive
 CC selection of a transformed cell and altering the level of expression of a
 CC polypeptide involved in post-transcriptional gene silencing in a host
 CC cell. The polynucleotides encoding polypeptides involved in gene
 CC silencing are useful in altering the post-transcriptional gene silencing
 CC pathway in plant cells. These are specifically useful in creating
 CC transgenic plants where the levels of polypeptides involved in post-
 CC transcriptional gene silencing are altered with respect to non-transgenic
 CC plants which would result in plants with an enhanced or deficient post-
 CC transcriptional gene silencing. The polynucleotides may also be used to
 CC manipulate transposon activity, meristem activity, plant architecture and
 CC development, proliferation of undifferentiated plant cells in culture for
 CC callus propagation, as probes for genetically and physically mapping the
 CC genes that they are part of and as markers for traits linked to those
 CC genes. The sequence presented is a polypeptide having post-
 CC transcriptional gene silencing activity.
 XX
 SQ Sequence 389 AA;
 Query Match 0.9%; Score 9; DB 6; Length 389;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 677 LKINAKLGG 685
 Db |||||
 89 LKINAKLGG 97
 RESULT 14
 AEC75730
 ID AEC75730 standard; protein; 389 AA.
 AC AEC75730;
 XX
 XX 17-NOV-2005 (first entry)
 DT
 XX Wheat Argonaute (AGO1) homolog FIS clone protein, SEQ ID NO: 50.
 DE
 XX Gene silencing; transgenic plant; cell proliferation; agriculture;
 KW crop improvement; Argonaute protein.
 KW
 XX
 OS Triticum aestivum.
 XX
 XX US2005204427-A1.
 FN
 XX
 PD 15-SEP-2005.
 XX
 XX 30-MAR-2005; 2005US-00093888.
 PF
 XX 18-JUN-2001; 2001US-0298973P.
 PR
 XX 17-JUN-2002; 2002US-00174363.
 XX
 PA (BUTL/) BUTLER K H.

PA (CAHO/) CAHOON R E.
 PA (HARV/) HARVELL L T.
 PA (RAFA/) RAFALSKI J A.
 PA (SAKA/) SAKAI H.
 XX
 PI Butler KH, Cahoon RE, Harvell LT, Rafalski JA, Sakai H;
 XX
 XX WPI; 2005-618219/63.
 DR N-PSDB; AEC75729.
 XX
 PT New polynucleotide encoding a polypeptide having post-transcriptional
 PT gene silencing activity, useful for manipulating plant architecture and
 PT development, or for propagating callus.
 XX
 XX Disclosure; SEQ ID NO 50; 107pp; English.
 XX
 CC The invention relates to a polynucleotide encoding a polypeptide having
 CC post-transcriptional gene silencing (PTGS) activity. The invention also
 CC relates to a method for producing transgenic plants. The sequences and
 CC method of the invention are useful for manipulating transposon activity,
 CC meristem activity, plant architecture and development or proliferation of
 CC undifferentiated plant cells in culture and for propagating callus. The
 CC present sequence is the wheat Argonaute (AGO1) homolog protein. This
 CC sequence is involved in PTGS activity.
 XX
 SQ Sequence 389 AA;
 Query Match 0.9%; Score 9; DB 9; Length 389;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 677 LKINAKLGG 685
 Db |||||
 89 LKINAKLGG 97
 RESULT 15
 ADQ80691
 ID ADQ80691 standard; protein; 850 AA.
 XX
 AC ADQ80691;
 XX
 XX 21-OCT-2004 (first entry)
 DT
 XX Arabidopsis thaliana TFL1-binding protein #3.
 DE
 XX TFL1-binding protein; plant growth control; biotechnology;
 KW fishing industry; screening.
 KW
 XX Arabidopsis thaliana.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 783..792
 FT /note= "Encoded by nucleotides 2347..2445"
 XX
 XX JP2004208572-A.
 FN
 XX 29-JUL-2004.
 PD
 XX 27-DEC-2002; 2002JP-00381220.
 PF
 XX 27-DEC-2002; 2002JP-00381220.
 PR
 XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
 PA
 XX WPI; 2004-538191/52.
 DR N-PSDB; ADQ80690.
 XX
 PT Novel AGO protein, comprising PA2 and Piwi domain, specifically binding
 PT to TFL1 protein that is derived from Arabidopsis thaliana, useful for
 PT controlling growth phase of plant.
 XX
 XX Claim 2; SEQ ID NO 6; 104pp; Japanese.

XX The invention comprises the amino acid and coding sequences of
CC Arabidopsis thaliana proteins which specifically bind to TFL1 protein.
CC The amino acid and coding sequences of the invention are useful for
CC screening TFL1 family proteins. The DNA and protein sequences of the
CC invention are also useful for controlling the growth phase of a plant,
CC and as a reagent for research in biotechnology and fishing industry. The
CC present amino acid sequence represents an Arabidopsis thaliana TFL1-
CC binding protein of the invention.
XX
SQ Sequence 850 AA;

Query Match 0.9%; Score 9; DB 8; Length 850;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 677 LKINAKLGG 685
|||
Db 569 LKINAKLGG 577

Search completed: July 5, 2006, 13:53:28
Job time : 199 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 5, 2006, 22:39:47 ; Search time 9528 Seconds
(without alignments)
8979.490 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 5349

Sequence: 1 MSSNFELEKGFVHSLDPE.....RHEMFLOTNVYKPGMSFA 1020

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abes/ABSSWEB_epool/US10645746/runat_05072006_110146_5084/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p
-USER=US10645746 @CGN 1_1_9528 @runat_05072006_110146_5084 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases :

EST:*

1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1189	22.2	796	2	BJ155598
C 2	1173	21.9	786	2	BJ143897
C 3	1088	20.3	633	2	BJ122122
C 4	1063	19.9	700	2	BJ150328

C 5	1045	19.5	710	2	BJ132853
C 6	1034	19.3	700	2	BJ155451
C 7	1033	19.3	635	2	BJ755060
C 8	1009	18.9	677	2	BJ779221
C 9	999	18.7	565	4	CB401772
C 10	995	18.6	583	2	BJ116268
C 11	983	18.4	554	4	CB398209
C 12	944	17.6	557	2	BJ105221
C 13	928	17.3	534	2	BJ127233
C 14	911	17.0	539	2	BJ127087
C 15	670	12.5	378	4	C62850
C 16	617	11.5	360	4	C60407
C 17	611	11.4	360	9	D35146
C 18	606	11.3	356	10	Z14900
C 19	605	11.3	360	4	C65335
C 20	605	11.3	360	9	D35838
C 21	590	11.0	373	4	C60787
C 22	584	10.9	370	1	AV191613
C 23	576	10.8	376	4	C60885
C 24	574	10.7	449	9	D27223
C 25	574	10.7	2465	14	AY417690
C 26	574	10.7	2465	14	AY417692
C 27	554	10.4	3778	6	AK035042
C 28	553	10.3	3845	6	AK030018
C 29	550.5	10.3	3160	6	AK154436
C 30	550.5	10.3	3160	6	AK171118
C 31	537.5	10.0	3127	6	AK083652
C 32	535	10.0	3606	6	AK080954
C 33	532	9.9	340	9	D27220
C 34	529.5	9.9	3123	13	CL969749
C 35	528	9.9	3509	6	BC073067
C 36	526	9.8	334	10	M89235
C 37	526	9.8	421	9	D27221
C 38	517.5	9.7	3188	6	CNS0AAU5
C 39	510	9.5	2955	13	CL965306
C 40	496.5	9.3	3168	13	CL963072
C 41	495	9.3	3222	13	CL969750
C 42	486	9.1	3336	13	CL967201
C 43	479	9.0	375	9	D32492
C 44	479	9.0	2522	14	DQ040105
C 45	476	8.9	295	9	D27222

ALIGNMENTS

RESULT 1	BJ155598/c	796 bp	mrna	linear	EST 24-JAN-2002
LOCUS	BJ155598	unpublished oligo-capped cdna library, C. elegans LI stage			
DEFINITION	Caenorhabditis elegans cdna clone yk1344b04 3', mRNA sequence.				
ACCESSION	BJ155598				
VERSION	BJ155598.1	GI:18323583			
KEYWORDS	EST.				
SOURCE	Caenorhabditis elegans				
ORGANISM	Caenorhabditis elegans				
REFERENCE	1 (bases 1 to 796)				
AUTHORS	Kohara.Y., Shin-i.T., Thierry-Mieg.J., Thierry-Mieg.D., Suzuki.Y. and Sugano.S.				
TITLE	A complementary view of the C.elegans genome				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.				
FEATURES	Location/Qualifiers				
source	1..796				
	/organism="Caenorhabditis elegans"				
	/mol_type="mRNA"				


```

Db      365  AAGGAATGAGCAAGTCTATATAAATGACCTACGACTTGCTTTCTCTGCT 306
Qy      959  ArgCysArgLysProIleSerLeuProValHisTyrAlaHisLeuSerCysGlu 978
Db      305  AGATGTGGAAGAACCATCTCGTTCGCTGTTCCGCTTCATTATGCTCATATCATGTGA 246
Qy      979  LysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnPro 998
Db      245  AAGCGAAGAGCTTTATCGNACTTACAGGACATTACATCGGTGACTATGCACGCCA 186
Qy      999  ArgThrArgHisGluMetGluHisPheLeuGlnThrAsnValLysTyrProGlyMetSer 1018
Db      185  CGGACTCGACAGAAATGGAACATTTCTCCAACTAAGCTGAAGTACCTCGGATGCTG 126
Qy      1019 PheAla 1020
Db      125  TTCGCA 120

RESULT 3
LOCUS   BJ122122
DEFINITION BJ122122 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION BJ122122
VERSION    BJ122122.1 GI:18282260
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS   Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
TITLE     A complementary view of the C.elegans genome
JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp.
          Location/Qualifiers
FEATURES
source
1..633
/mol_type="mRNA"
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1284h06"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

ORIGIN
Alignment Scores:
Pred. No.: 1.9e-113 Length: 633
Score: 1088.00 Matches: 209
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.1% Mismatches: 1
Query Match: 20.3% Indels: 0
DB: 2 Gaps: 0

US-10-645-746-3 (1-1020) x BJ122122 (1-633)

Qy      453  LeuLysMetLeuLysGluLeuAspPheSerSerGluLeuAsnPheValGluArgPhe 472
Db      1  CTAAAAATGCTGAAGAAATGGATTTCTCTTCGAAGAGCTAAATTTGTGGAAGATTT 60
Qy      473  GlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMet 492
Db      61  GGATTATGCTCCAACTTCAGATGATCGAATGTCAGGAAGAGGTTTTTGAAGAGCCATG 120

```

```

Qy      493  LeuValAsnSerValAsnGluGlnIleLysMetThrProValIleArgGlyPheGlnGlu 512
Db      121  CTTGTGAATAGTGTAAATGAACAAATTAATAATGACACAGTGATTCGTGGATTTCAAGA 180
Qy      513  LysGlnLeuAsnValValProGluLysGluLeuCysCysAlaValPheValValAsnGlu 532
Db      181  AAACAATTTGAATGTGGTTCCGAAAAAGAACTTTTGCTGTGTCTGTTTTGTAGCCCAACGA 240
Qy      533  ThrAlaGlyAsnProCysLeuGluGluAsnAspValValLysPheTyrThrGluLeuIle 552
Db      241  ACAGCGGGAATTCATGCTTAGAAGACAGACGTTGTTAAGTTCTACACCGAATAATT 300
Qy      553  GlyGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSer 572
Db      301  GGTGGTTTGCAGATTCGCTGGAATACGAATTTGGTGCCAATGAAACACAGAGGCGCAATCT 360
Qy      573  IleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThr 592
Db      361  ATTATGTATACGACGCGACGAAAAATGAATATGCTTCTTACAAAAATTTGTACATAATACC 420
Qy      593  GlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuPro 612
Db      421  GGAATCGGTAGATTTGAATAGCGCAACAGAGCGGAAGATATGTTTGACGCTCTTCCC 480
Qy      613  AspLysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyr 632
Db      481  GATAAAGAACACCAAGTCTTAATGTTTCATTATCATTTCCAAACGACAACTGAATGCTTAC 540
Qy      633  GlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSer 652
Db      541  GGTTTTGTGAAACATTATTTCGATCATCACCATCGGTGTAGCTAATCAGCATATTTACTTCT 600
Qy      653  GluThrValThrLysAlaLeuAlaSerLeuArg 663
Db      601  GAAACAGTCAAAAAGCTTTGGCATCTACTAAGG 633

```

BJ150328 700 bp mRNA linear EST 24-JAN-2002
 BJ150328 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1284h06 3', mRNA sequence.

BJ150328
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

A complementary view of the C.elegans genome
 Unpublished (2002)
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source

1..700
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1284h06"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

ORIGIN

Alignment Scores:

Pred. No.: 1.64e-110 Length: 700
 Score: 1063.00 Matches: 200
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 19.9% Indels: 0
 DB: 2 Gaps: 0

US-10-645-746-3 (1-1020) x BJ150328 (1-700)

QY 821 SerGluValLysGluPheMetSerGluArgAspGlyGluAspProGluProLysThr 840
 DB 699 AGCGAAGTAAACAATTCATGTCGGAACGGATGGAGAAGATCCAGAGCCGAAGTACAG 640
 QY 841 PheIleValIleGlnLysArgHisAenThrArgLysLeuArgMetGluLysAspLys 860
 DB 639 TTCATTGTGATTCAGAAAGACACATACACGATTGCTTCGAAGATGGAAAAAGATAAG 580
 QY 861 ProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys 880
 DB 579 CCAGTGGTCAATAAAGATCTTACTCTGCTGGAACAGATGTCGCTGTTGCTGCTGTTAAA 520
 QY 881 GlnTTPGluGluAspMetLysGluSerLysGluThrGlyLleValAsnProSerSerGly 900
 DB 519 CAATGGGAGGAGATATGAAGAAGACCAAGAAACTGGAAATTTGTGAACCCATCATCCGGA 460
 QY 901 ThrThrValAspLysLysLeuValSerLysThrLysPheAspPhePheLeuAlaSerHis 920
 DB 459 ACAACTGTGGATAAATTCATGTTCCGAATAACAATTCGATTTTCTTGGCGTCTCAT 400
 QY 921 HisGlyValLeuGlyThrSerArgProGlyHisThrValMetLysAspLysGly 940
 DB 399 CATGGTGTCTTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACGATAAAGGA 340
 QY 941 MetSerGlnAspGluValLysLysMetThrLysGlyLysLeuAlaPheLeuSerAlaArgCys 960
 DB 339 ATGACCCAAAGATGAAAGTCTATAAATGACCTACCGACTTGTCTTCTCTGCTAGATGT 280
 QY 961 ArgLysProIleSerLeuProValProValHisThrAlaHisLeuSerCysGluLysAla 980
 DB 279 CGAAAAACCATCTCGTGTGCTTCTCGGTTTCATTATGCTCATTTATCATGTCAAAAAGCG 220
 QY 981 LysGluLeuThrArgThrLysGluHisThrLysLysGlyAspThrAlaGlnProArgThr 1000
 DB 219 AAAGAGCTTTATCGAACTTACAAAGAAACATTACATCGGTGACTATGCACAGCCAGGACT 160
 QY 1001 ArgHisGluMetGluHisPheLeuGlnThrAsnValLysThrProGlyMetSerPheAla 1020
 DB 159 CGACACGAATGGAACATTTTCTCCAACTAACGTGAAAGTACCTTGGAAATGTGCTTCGCA 100

RESULT 5

BJ132853/c

LOCUS

BJ132853 710 bp mRNA linear EST 23-JAN-2002
 BJ132853 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1070f12 3', mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Caenorhabditis elegans

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 710)

Kohara,Y., Shin-I., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.

and Sugano,S.

A complementary view of the C.elegans genome

Unpublished (2002)

Contact: Tadasi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

FEATURES

source

1..710
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1070f12"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"

ORIGIN

Alignment Scores:
 Pred. No.: 1.95e-108 Length: 710
 Score: 1045.00 Matches: 196
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 19.5% Indels: 0
 DB: 2 Gaps: 0

US-10-645-746-3 (1-1020) x BJ132853 (1-710)

QY 825 GlnPheMetSerGluArgAspGlyGluAspProGluProLysThrThrPheIleValIle 844
 DB 705 CAATTCATGTCGGAACGGATGGAGAAGATCCAGAGCCGAAGTACACGTTTCATTTGTGATT 646
 QY 845 GlnLysArgHisAenThrArgLeuLeuArgMetGluLysAspLysProValValAen 864
 DB 645 CAGAAAAGACACAATACACGATTGCTTCGAAGAATGGAAAAAGATAAGCCAGTGGTCAAT 586
 QY 865 LysAspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGlu 884
 DB 585 AAAGATCTTACTCTGCTGAAACAGATGTCGCTGTTGCTGCTGCATCAACAAATGGGAGAG 526
 QY 885 AspMetLysGluSerLysGluThrGlyLleValAsnProSerSerGlyThrThrValAsp 904
 DB 525 GATATGAAGAAAGCAAAAGAACTGGAATTTGTAACCCATCATCCGGAACAACATGTGGAT 466
 QY 905 LysLeuIleValSerLysThrLysPhePhePheLeuAlaSerHisGlyValLeu 924
 DB 465 AAACCTATCGTTTCGAAATACAAATTCGATTTTCTTGGCATCTCATCATGTTGCTCTT 406
 QY 925 GlyThrSerArgProGlyHisThrValMetThrAspLysGlyMetSerGlnAsp 944
 DB 405 GGTACATCTCGTCCAGGACATTACACTGTTATGATGACGATAAAGGAATGAGCCAAGAT 346
 QY 945 GluValThrLysMetThrThrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIle 964
 DB 345 GAAGTCTATAAATGACCTACGAGCTTCTCTCTCTGCTAGATGTGCAAAACCCATC 286
 QY 965 SerLeuProValProValHisThrAlaHisLeuSerCysGluLysAlaLysGluLeuTyr 984
 DB 285 TCGTTGCTGTTCCGTTTCATTTATGCTATTTATGATGTAAGAAAGCGAAGGCTTTAT 226
 QY 985 ArgThrThrLysGluHisThrLysGlyAspThrAlaGlnProArgThrArgHisGluMet 1004
 DB 225 CGAAGTCTAACAGGAACATTACATCGTGACTATGACGACCCAGCCAGCTCGACACCAAGATG 166
 QY 1005 GluHisPheLeuGlnThrAsnValLysThrProGlyMetSerPheAla 1020
 DB 165 GAACATTTTCTCCAAACTAACGTGAAGTACCTTGGAAATGTGCTTCGCA 118

RESULT 6

BJ155451/c

LOCUS

DEFINITION

ACCESSION

UNPUBLISHED

Caenorhabditis elegans cDNA clone yk1342d11 3', mRNA sequence.

BJ155451

700 bp mRNA linear EST 24-JAN-2002

BJ155451 unpublished oligo-capped cDNA library, C. elegans L1 stage

Caenorhabditis elegans cDNA clone yk1342d11 3', mRNA sequence.

```

VERSION      BJ155451.1  GI:18323436
KEYWORDS     EST.
SOURCE       Caenorhabditis elegans
ORGANISM     Caenorhabditis elegans
REFERENCE    1 (bases 1 to 700)
AUTHORS      Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE        A complementary view of the C.elegans genome
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp.
              Location/Qualifiers
FEATURES     source
             1..700
                /organism="Caenorhabditis elegans"
                /mol_type="mRNA"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="yk1342d11"
                /sex="hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L1"
                /clone_lib="unpublished oligo-capped cDNA library, C.
                elegans L1 stage"
ORIGIN
Alignment Scores:
Pred. No.:      3,49e-107      Length:      700
Score:          1034.00      Matches:    194
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    19.3%      Indels:     0
DB:             2      Gaps:         0

US-10-645-746-3 (1-1020) x BJ155451 (1-700)

QY  827 MetSerGluArgAspGlyGluAspProGluProLysThrPheIleValIleGlnLys 846
Db  699 ATGTCGGAACGGATGGAGAAGATCCAGACCCGAGTACAGTTCATTTGGATTCAGAA 640

QY  847 AtgHiaAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLysAsp 866
Db  639 AGACACAATACAGATGCTTCGAGAGATGGAAAAGATAGCCAGTGGTCATTAAGAT 580

QY  867 LeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMet 886
Db  579 CTTACTCCTGCTGAAACAGATGTCGCTGTTGCTGTTAAACAATGGGAGGAGGATATG 520

QY  887 LysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAlaLysLeu 906
Db  519 AAAGAAAGCAAGAAATGGAATTTGTAACCCCATCATCCGGAACAATGTTGGATAAAT 460

QY  907 IleValSerLysThrLysPheAspPhePheLeuAlaSerHisGlyValLeuGlyThr 926
Db  459 ATGTTTCGAAATACAAATTCGATTTTTCCTGGCATCTCATCATGGTGTCTTGTGACA 400

QY  927 SerArgProGlyHisThrValMetTyrAspAspLysGlyMetSerGlnAspGluVal 946
Db  399 TCTCGTCCAGGACATACACTGTTATGTATGACGATAAAGGAATGAGCAAGATGAATC 340

QY  947 TyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeu 966
Db  339 TATAAATGACCTACCGACTGCTTTCTCTCGCTAGATGTCGAAACCCATCTCGTTG 280

QY  967 ProValProValHisThrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThr 986
Db  279 CTTGTTCCGGTTCATTTATGCTCATTTATCATGTGAAAGAGCGAAGAGCTTTATCGA 220

```

```

QY  987 TyrLysGluHisThrIleGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHis 1006
Db  219 TACAGGACACATTACATCGTGACTATGCACAGCCAGGACTGCACACGAAATGGAACAT 160

QY  1007 PheLeuGlnThrAsnValLysTyrProGlyMetSerPheAla 1020
Db  159 TTCTCCAAACTAACGTGAAGTACCCCTGGAAATGTCGTTCCGA 118

RESULT 7
LOCUS    BJ755060
DEFINITION BJ755060 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk1409d07 5', mRNA sequence.
ACCESSION BJ755060
VERSION    BJ755060.1  GI:47594822
KEYWORDS   EST.
SOURCE     Caenorhabditis elegans
ORGANISM   Caenorhabditis elegans
REFERENCE  1 (bases 1 to 635)
AUTHORS    Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE      A complementary view of the C.elegans genome
JOURNAL    Unpublished (2002)
COMMENT    Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp.
              Location/Qualifiers
FEATURES   source
           1..635
              /organism="Caenorhabditis elegans"
              /mol_type="mRNA"
              /strain="N2"
              /db_xref="taxon:6239"
              /clone="yk1409d07"
              /sex="Hermaphrodite"
              /tissue_type="whole animal"
              /dev_stage="varied"
              /clone_lib="unpublished oligo-capped cDNA library"
ORIGIN
Alignment Scores:
Pred. No.:      3,89e-107      Length:      635
Score:          1033.00      Matches:    194
Percent Similarity: 99.5%      Conservative: 0
Best Local Similarity: 99.5%      Mismatches: 1
Query Match:    19.3%      Indels:     0
DB:             2      Gaps:         0

US-10-645-746-3 (1-1020) x BJ755060 (1-635)

QY  1 MetSerSerAsnPheProGluLeuLysGlyPheTyrArgHisSerLeuAspProGlu 20
Db  51  ATGTCCTCGAATTTTCCCGAATTTGGAAAAAGGATTTTATCGTCATCTCTCGATCCGAG 110

QY  21 MetLysThrPheAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysVal 40
Db  111 ATGAAATGGCTTCGCGAGGCCCATCTGGTAAATGCGACGCAAAATTTCTATGAGAAAGTA 170

QY  41 LeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGlu 60
Db  171 CTTCTTTTGGTAAATTTGGTTCAAGTTCTCAGCAAAATTTACGATCGGAAATCTACGAG 230

QY  61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
Db  281 TATGAAGTCAAAATGCAAAAGGAAGTATTGAATAGAAACCCAGGAAACCTTTCCCAAAA 290

QY  81 LysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGlu 100

```

```

Db      291  AGACAGAAATTCCTCCGATCGTGCAGAACTCTTCTGCTACATCTTTCGGCATGAG 350
QY      101  LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
Db      351  AAGAGCAGACAGATTTTATCTCGAAGACTATGTTTTTGTGATGAAAGGACACTGTTTAT 410
QY      121  SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValValLys 140
Db      411  AGTGTGTCGACTGCACTGTGCATCAAAATGCTGGTTTCGGAGAAAGTAGTAAAA 470
QY      141  LysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTyrThrMetIle 160
Db      471  AGGATTCGGAGAAAAGATGAAAGGATTTGGAGAAAATACTTATACATGATA 530
QY      161  LeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAspGlu 180
Db      531  CTTACCTATCGTAAAAATTTTCACCTGAACTTTAGTCGAGAAAAATCCGGAAAAAGACGAA 590
QY      181  GluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln 195
Db      591  GAAGCGAATCGGAGTTTACAAATTCCTGAAGAATGTTATGACCCAG 635

```

```

RESULT 8
BJ779221/c
LOCUS   BJ779221 677 bp mRNA linear EST 25-MAY-2004
DEFINITION BJ779221 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone Yk1409d07 3', mRNA sequence.
ACCESSION BJ779221
VERSION   BJ779221.1 GI:47654855
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans

```

```

REFERENCE 1 (bases 1 to 677)
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tschini@genes.nig.ac.jp.
Location/Qualifiers
1. .677
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1409d07"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 2 44e-104 Length: 677
Score: 1009.00 Matches: 189
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 18.9% Indels: 0
DB: 2 Gaps: 0
US-10-645-746-3 (1-1020) x BJ779221 (1-677)

```

```

QY      832 GlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArg 851
Db      675  GGAGAAGATCCAGAGCCGGAAGTACACGTTTCATTGTGATTTCAGAAAAGACACATACAGCA 616

```

```

QY      852  LeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGlu 871
Db      615  TTGCTTCGAAGAATCGAAAAGATAGCCAGTGGTCAATAAAAGATCTTTACTCTGCTGAA 556
QY      872  ThrAspValAlaValAlaValLysGlnTspGluGluAspMetLysGluSerLysGlu 891
Db      555  ACAGATGTCGCTGTTGCTGCTTTAAACATGGGAGGAGGATATGAAAGAAAGCAAGAA 496
QY      892  ThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyr 911
Db      495  ACTCGAAATTGTGAACCCATCATCCGGAACAACATGTGGATAAACTTATCGTTTCGAAATAC 436
QY      912  LysPheAspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHis 931
Db      435  AAATTCGATTTTTTCTTGGCATCTCATCATGGTGTCTTGGTACATCTCGTCAGGACAT 376
QY      932  TyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyr 951
Db      375  TACACTGTTATGATGACGATAAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTAC 316
QY      952  GlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHis 971
Db      315  GGACTTGGCTTTCTCTCTAGATGTGAAAACCCATCTCGTTGCTTCTCGGTTTCAT 256
QY      972  TyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyr 991
Db      255  TATGCTCATTTATCATGTGAAAAGCGAAAGAGCTTTATCGAACTTTACAAAGAACATTAC 196
QY      992  IleGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThrAsn 1011
Db      195  ATCGTGACTATGCACAGCCAGGACTCGACACGAATGGAACATTTTCTCCAAACTAAC 136
QY      1012  ValLysTyrProGlyMetSerPheAla 1020
Db      135  GTGAAGTACCCTCGGAATGTCGTTTCGA 109

```

```

RESULT 9
CB401772
LOCUS   CB401772 565 bp mRNA linear EST 15-MAY-2003
DEFINITION OSTF199B4_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION CB401772
VERSION   CB401772.1 GI:30743499
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans

```

```

REFERENCE 1 (bases 1 to 565)
AUTHORS Rebol,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M.,
Armstrong,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T.,
Hudson,J.R., Hartley,J.B., Braech,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevet,E., Papasotiropoulos,V.,
Tolias,P.P., Ptacek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.B. and Vidal,M.
C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
Nat. Genet. (2003) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hill@dfci.harvard.edu or
marc-vidal@dfci.harvard.edu
POLYA=No.
Location/Qualifiers
1. .565

```

```

JOURNAL Nat. Genet. (2003) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hill@dfci.harvard.edu or
marc-vidal@dfci.harvard.edu
POLYA=No.
Location/Qualifiers
1. .565

```

```

FEATURES
source
1. .565

```



```

Db      481 TCTCTCTGGATATCTTCTCTCTTAATGTGCGACCCCGAGTGTGTAACGATGATGACGA 540
Qy      301 LysAppLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
Db      541 AAAGATCTTAAACAAACTGATGGCGGGAATAATGACAATC 582

RESULT 11
CB398209/c
LOCUS   CB398209_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA linear EST 15-MAY-2003
DEFINITION
ACCESSION CB398209
VERSION   CB398209.1 GI:30739936
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE
AUTHORS Rebol,J., Vaglio,P., Rual,J.P., Lamesch,P., Martinez,M.,
Armstrong,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T.,
Hudson,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevet,E., Papasotiropoulos,V.,
Tollas,P.P., Placek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
Nat. Genet. (2003) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hilledfci.harvard.edu or
marc_vidal@dfci.harvard.edu
POLYA=No.
FEATURES             Location/Qualifiers
     source           1..554
                     /organism="Caenorhabditis elegans"
                     /mol_type="mRNA"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /sex="Hermaphrodite and male"
                     /tissue_type="whole animal"
                     /dev_stage="mixed stage"
                     /clone_lib="AD-wrmcDNA"
                     /note="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pFC86"
ORIGIN
Alignment Scores:
Pred. No.:      1.71e-101      Length:      554
Score:          983.00        Matches:    184
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    18.4%          Indels:     0
DB:             4              Gaps:      0

US-10-645-746-3 (1-1020) x CB398209 (1-554)

Qy      831 AspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisLenThr 850
Db      553 GATGGAGAGATCCAGAGCCGGAAGTACACGTTCATTGTGATTTCAGAAAGACACATACA 494
Qy      851 ArgLeuLeuArgMetGluLysAspLysProValValAsnLysAspLeuThrProAla 870

```

```

Db      493 CGATTGCTTCGAAGATGGAATAAGCAGTGGTCAATAAGATCTTACTCTCTGCT 434
Qy      871 GluThrAspValAlaValAlaValLysGlnTTPGluGluAspMetLysGluSerLys 890
Db      433 GAAACAGATGTCGCTGTTTAAACAATGGGAGGAGATATGAAGAAGCAAA 374
Qy      891 GluThrGlyIleValAsnProSerSerGlyThrThrValAspLysValSerLys 910
Db      373 GAAACTGGAATTTGTGAACCCATCATCCGGAACACTGTGGATAAACTTATCGTTTCGAAA 314
Qy      911 TyrLysPheAspPhePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGly 930
Db      313 TACAATTCGATTTTCTTGGCATCTCATCATGTGTGCTTGGTACATCTCGTCCAGCA 254
Qy      931 HisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThr 950
Db      253 CATTAACCTGTTATGTATGACGATTAAGGAATGAGCAAGATGAAGTCTTATAAAATGACC 194
Qy      951 TyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProVal 970
Db      193 TACGACACTGCTTTTCTCTCTGTAGATGTGGAACCCATCTCGTTGCTGCTGTTCCGGT 134
Qy      971 HisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db      133 CATTAATGCTCATTTATCATGTGAAAAGCGAAAGAGCTTTATCGAACTTACAAGGAACAT 74
Qy      991 TyrIleGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThr 1010
Db      73 TACATCGGTGACTATGCACGACCGACGACGAAATGGAACATTTTCTCCAAACT 14
Qy      1011 AsnValLysTyr 1014
Db      13 AACGTGAAGTAC 2

RESULT 12
BJ105221
LOCUS   BJ105221
DEFINITION
ACCESSION BJ105221
VERSION   BJ105221.1 GI:18247891
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 557)
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE    A complementary view of the C.elegans genome
JOURNAL  Unpublished (2002)
COMMENT  Contact: Tadashi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source           1..557
                     /organism="Caenorhabditis elegans"
                     /mol_type="mRNA"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone="yk1070f12"
                     /sex="hermaphrodite"
                     /tissue_type="whole animal"
                     /dev_stage="L1"
                     /clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
ORIGIN

```


Alignment Scores: 5.13e-97 Length: 557
 Pred. No.: 944.00 Matches: 184
 Score: 944.00 Conservative: 1
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 99.5% Indels: 0
 Query Match: 17.6% Gaps: 0
 DB: 2

US-10-645-746-3 (1-1020) x BJ105221 (1-557)

QY 675 IleAlaLeuYrIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTrrSerGlu 694
 Db 1 ATTGCATTGAATCAACGCGAATATTAGGAGTATTACCCAGAGCTTGACTGCTCAGAA 60

QY 695 IleAlaGluIleSerProGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 714
 Db 61 ATTGCAGATAATATCACCAGAGAAAGAAAGACGGAACCAATGCCATTAATATGAT 120

QY 715 ValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAla 734
 Db 121 GTTGGAAATTGATGTAATCTATCCAACTCTCCACAGTGGAAATTGATTATCTATAGCGGCT 180

QY 735 ValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGlu 754
 Db 181 GTAGTAGCGAGTATCAATCCAGGTGGAACTATCTATCGAATATGATTGTGACTCAGAA 240

QY 755 GluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu 774
 Db 241 GAATGTGTCGCGTGAGCGTGCAGTGGCTCATGCGCGGAAAGAACAGATATTTTGGAA 300

QY 775 AlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAspAsnArgAlaProAla 794
 Db 301 GCAAGTTCGTGAATTCGAGAAATTCGAGAAATTCGAGAAATTCGAGAAATTCGAGAA 360

QY 795 HisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAsp 814
 Db 361 CATATTGTAGTCTATCGAGACGGAGTTAGCGATTCGAGATGCTAGCTGTAGTCATGAT 420

QY 815 GluLeuArgSerLeuLysSerGluValLysGluPheMetSerGluArgAspGlyGluAsp 834
 Db 421 GAGCTTCGATCTTTAAGAACGGAAGTAAACAAATTCATGTCCGAAACGGGATCGAGAAT 480

QY 835 ProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArg 854
 Db 481 CCGAGCCGAGTACAGCTTCATGTGATTGATTGAGAAAGACAAATACAGATTGCTTCGA 540

QY 855 ArgMetGluLysAsp 859
 Db 541 AGAATCGAAAAAGAT 555

RESULT 13
 BJ127233
 LOCUS
 DEFINITION
 Caenorhabditis elegans cDNA library, C. elegans L1 stage
 BJ127233
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

534 bp mRNA linear EST 23-JAN-2002
 BJ127233 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1342d11 5', mRNA sequence.

BJ127233.1 GI:18287390
 EST.
 Caenorhabditis elegans
 Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 534)
 Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
 and Sugano, S.
 A complementary view of the C. elegans genome
 Unpublished (2002)
 Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855

Email: tchini@genes.nig.ac.jp.
 Location/Qualifiers
 1. 534
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1342d11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

FEATURES
 source

ORIGIN
 Alignment Scores: 3.28e-95 Length: 534
 Pred. No.: 928.00 Matches: 177
 Score: 928.00 Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 17.3% Gaps: 0
 DB: 2

US-10-645-746-3 (1-1020) x BJ127233 (1-534)

QY 363 LysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPhe 382
 Db 3 AAAGTCACTGGTAAATCGGACAGAGGAAGAAATGCAAAAAAGTAGTACGATACCTACATGTTTC 62

QY 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
 Db 63 AAAATCTATGAGGAAACAAAAAGTTTCATTTGAGTTTCCCCACCTACCACTAGTCAAGTT 122

QY 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422
 Db 123 AAAAGTGGACAAAGATACGCTGTACCAATGGAACATCTTGAAGTTTCATGAGAGCA 182

QY 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
 Db 183 CAAAGATACAAAGATCGAATTTGATCTGGTATGCAAGCAAGTTTCTAAAGCGAGCTACA 242

QY 443 ArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLysAspPheSer 462
 Db 243 CGAAAAACCTCACACTACAAAGAAAAATACCTTAAAAATGCTGAAAGAAATTTGGATTCTCT 302

QY 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu 482
 Db 303 TCTGAAGAGCTAAATTTGTTTGAAGATTTGGATTATGCTCCAAACTTCAGATGATCGAA 362

QY 483 CysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLys 502
 Db 363 TGTCAGGAAAGGTTTTGAAAGAGCAATGCTTGTGAATAGTGTAAATGAACAAATTTAA 422

QY 503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGlu 522
 Db 423 ATGACACCACTGATTCGTGGATTTCAAGAAAAACAATTTGAATGTTGGTTCGCGAAAAAGAA 482

QY 523 LeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeu 539
 Db 483 CTTTGTCTGTCTGTTTTTTTGTAGTCAAGAAACAGCGGAAATCCATGCTTA 533

RESULT 14
 BJ127087
 LOCUS
 DEFINITION
 Caenorhabditis elegans cDNA library, C. elegans L1 stage
 BJ127087
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

539 bp mRNA linear EST 23-JAN-2002
 BJ127087 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1342d11 5', mRNA sequence.

BJ127087.1 GI:18287244
 EST.
 Caenorhabditis elegans
 Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 539)
AUTHORS Kohara,Y., Shin-I.T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
Location/Qualifiers

Source
1..539
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1342d11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

ORIGIN

Alignment Scores:
Pred. No.: 2,97e-93 Length: 539
Score: 911.00 Matches: 177
Percent Similarity: 98.3% Conservative: 0
Best Local Similarity: 98.3% Mismatches: 2
Query Match: 17.0% Indels: 1
DB: 2 Gaps: 0

US-10-645-746-3 (1-1020) x BJ127087 (1-539)

QY 306 LysLeuMetAlaGlyLysMetThrIleArgGlnAlaAaArgProArgIleArgGlnLeu 325
DB 2 AAACGTATCGGGGAAAAATGACATACAGCAAGCCGCGGCCCAAGATTTCACAAATTA 61
QY 326 LeuGluAsnLeuLysLeuLysCysAlaGluValTTPAspAsnGluMetSerArgLeuThr 345
DB 62 TTGGAAATTTGAAGCTGAATTCGCGAGAAGTTTGGGATAACGAAATGTCGAGATTGACA 121
QY 346 GluAaGHisLeuThPhelLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThr 365
DB 122 GAACGACATCTGACATTTCTAGATTTGTGCGANGAAACCTCTCTGTTTATAAAGTCAT 181
QY 366 GlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyr 385
DB 182 GGTAAATCGACAGAGAGAGAAATGCANAAAGTAGCTACTACATTGTTCAAAATCTAT 241
QY 386 GluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGly 405
DB 242 GAGGAAAAACAAAAGTTTCAATGAGTTTCCCACTACCACCTANTCAAAGTTAAAGTGA 301
QY 406 AlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyr 425
DB 302 GCAAAAGAAATACGCTGTACCAATGGAAACATCTTGAAGTTTCATGAGAAAGCCACAAAGATAC 361
QY 426 LysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysPro 445
DB 362 AGAATCGAATTGATCTGGTATGCAAGCAAGTTTCTAAGCGAGCTACACGAAACCT 421
QY 446 HisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGlu 465
DB 422 CACGACTACAAAGAAAATACCCCTAAATAATGCTGAAAGAAATTCGATTTCTCTTCTGAAGAG 481
QY 466 LeuAsnPheValClnArgPheGlyLeuCysSerLysLeuGlnMetIle-GluCysPro 484
DB 482 CTAATTTTGTGAAAGATTGGATTATGCTCCAAACITTCAGATGATCCGAATGTCGA 539

RESULT 15

C62850

LOCUS
DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 22, 1997 this sequence version replaced gi:2421555.

Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

Location/Qualifiers

Source
1..378
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk296b10"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="Yuji Kohara unpublished cDNA"

ORIGIN

Alignment Scores:
Pred. No.: 7,51e-66 Length: 378
Score: 670.00 Matches: 122
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 2
Query Match: 12.5% Indels: 0
DB: 4 Gaps: 0

US-10-645-746-3 (1-1020) x C62850 (1-378)

QY 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
DB 7 ATGCTCTCGAATTTTCCGAAATTGGAAGGAGATTATCGTCATCTCTCGATCCGGAG 66
QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysLysVal 40
DB 67 ATGAAATGCTTTCGAGGCCCCACTGGTANATCGACGCGCAAAATTTATGAGAAAGATGA 126
QY 41 LeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGlu 60
DB 127 CTTCTTTTGGTAAATTTGGTTCAAGTTCTCCACCAAAATTTACATCGGGAATACTACGAG 186
QY 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
DB 187 TATGAAGTGAAATTCACAAAGGAAGTNTTGAATAGAAAACCAAGGAAAACCTTTCCCAAAA 246
QY 81 LysThrGluIleProIleProAspArgAlaLysLeuPheTrpGluHisLeuArgHisGlu 100
DB 247 AAGACAGAAATTTCCAAATTTCCCATTTCCCATTTCTCTGCAACATCTTCGGCATGAG 306
QY 101 LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
DB 307 AAGAGGACAGACAGATTTTATTTCTCGAGACTATGTTTTTGTATGAAANGACACTGTTTAT 366
QY 121 SerValCysArg 124
|||||

Db 367 AGTGTTCGCA 378

Search completed: July 6, 2006, 04:44:16
Job time : 9539 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 5, 2006, 22:12:21 ; Search time 1292 Seconds
(without alignments)
8256.610 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 M5SNFPELEKGFGRHSLDPE.....RHEMEHFLQTNVKYPGMSFA 1020

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
-Q-/abss/ABSSWEB	Ygapop 10.0 , Ygapext 0.5
-UNITS-bits	Fgapop 6.0 , Fgapext 7.0
-DOCALIGN=200	Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q-/abss/ABSSWEB spool/US10645746/runat_05072006_110142_5038/app query.fasta.1
-DB=N Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p
-USER=US10645746 @CGN 1.1 1147 @runat_05072006_110142_5038 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 8:*

1:	Geneseqn1990s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*
14:	Geneseqn2005s:*
15:	Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5349	100.0	3227	4	Aas03283
2	4804.5	89.8	3709	4	Aas03282 C. elegans
3	504.5	11.3	2960	13	Adq80698 Arabidops

4	577	10.8	2571	12	ADJ94703	Adj94703 Human euk
5	577	10.8	7478	10	ADB81486	ADB81486 DNA of th
6	577	10.8	7478	14	AEC01623	Aec01623 Human EIF
7	575.5	10.8	2910	9	ACD98411	Acd98411 A. thalia
8	575.5	10.8	2910	10	ADC46702	Adc46702 Thaleares
9	575.5	10.8	2910	12	ADI43620	Adi43620 Plant tra
10	575.5	10.8	2910	12	ADO02338	Ado02338 Thaleares
11	566	10.6	2580	12	ADJ94704	Adj94704 Human euk
12	566	10.6	3011	6	ABQ99304	Abq99304 Human cod
13	566	10.6	3580	6	ABS78717	Abs78717 Human cdn
14	565	10.6	3996	4	AHL14510	Ahl14510 Human cdn
15	564.5	10.6	2955	6	AHS1392	Ahs1392 CDNA enco
16	564.5	10.6	3486	4	ABL20457	Ab120457 Drosophil
17	564.5	10.6	3586	4	ABL20459	Ab120459 Drosophil
18	564.5	10.6	3743	4	ABL11513	Ab111513 Drosophil
19	556.5	10.4	2568	12	ADJ94706	Adj94706 Human euk
20	556.5	10.4	2914	4	AAK53425	Aak53425 Human pol
21	556.5	10.4	2914	8	ABX13625	Abx13625 Human cyt
22	556.5	10.4	3050	4	AAH16058	Aah16058 Human cdn
23	556.5	10.4	3050	12	ADH77114	Adh77114 Human PAZ
24	555.5	10.4	3736	8	ABX34752	Abx34752 Human mdd
25	555	10.4	2772	12	ADJ94705	Adj94705 Human euk
26	550.5	10.3	3307	14	ADM42124	Adm42124 Human euk
27	550.5	10.3	3325	8	ADM55860	Adm55860 Human nuc
28	530.5	9.9	2815	13	ADP55695	Adp55695 Human PRO
29	523.5	9.8	2960	13	ADQ80704	Adq80704 Arabidops
30	523.5	9.8	2973	12	ADO61720	Ado61720 Transcrip
31	519.5	9.7	3008	12	ADO63642	Ado63642 Transcrip
32	519.5	9.7	8465	4	ABL20458	Ab120458 Drosophil
33	519.5	9.7	10149	4	ABL20456	Ab120456 Drosophil
34	519.5	9.7	12558	4	ABL11512	Ab111512 Drosophil
35	519	9.7	3120	13	ADQ80694	Adq80694 Arabidops
36	519	9.7	3147	12	ADI43622	Adi43622 Plant tra
37	519	9.7	3326	3	AAC47079	Aac47079 Arabidops
38	515	9.6	3071	14	ADY65087	Ady65087 S. manson
39	510.5	9.5	3040	13	ADQ80700	Adq80700 Arabidops
40	509	9.5	3195	3	AAC50992	Aac50992 Arabidops
41	501	9.4	2960	13	ADQ80696	Adq80696 Arabidops
42	501	9.4	2967	3	AAC50603	Aac50603 Arabidops
43	501	9.4	2967	6	ABZ12627	Abz12627 Arabidops
44	496.5	9.3	3718	13	ADX45492	Adx45492 Plant ful
45	495	9.3	2042	9	ADB17457	Adb17457 Corn post

ALIGNMENTS

RESULT 1

AAS03283
ID AAS03283 standard; cDNA; 3227 BP.

XX AAS03283;

XX 07-SRP-2001 (first entry)

XX C. elegans cDNA encoding RNA interference pathway protein RDE-1.

XX RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;
XX double-stranded RNA-dependent gene silencing; ss.

XX Caenorhabditis elegans.

Key	Location/Qualifiers
5'UTR	1..20
FT	/*tag= a
FT	21..3083
FT	/*tag= b
FT	/product= "RDE-1"

XX WO200129058-A1.

XX 26-APR-2001.

XX 13-OCT-2000; 2000WO-US028470.

XX 15-OCT-1999; 99US-0159776P.
 PR 30-MAR-2006; 2000US-0193218P.
 XX (UYNA-) UNIV MASSACHUSETTS.
 PA (CARN-) CARNegie INST WASHINGTON.
 XX Mello CC, Fire A, Tabara H, Grishok A;
 PI WPI; 2001-316239/33.
 DR P-PSDB; AAU01856.
 XX Novel RNA interference pathway genes and their protein products involved
 PT in mediation of genetic interference, useful for modulating and studying
 PT regulation of RNA interference pathway.
 XX Claim 1; Fig 6; 76pp; English.
 PS The sequence encodes the RNA interference (RNAi) pathway protein RDE-1.
 CC RDE-1 and RDE-4 are involved in the pathway mediating double-stranded
 CC RNA-dependent gene silencing (genetic interference) RDE-1 and RDE-4
 CC protein is useful for preparing an RNAi agent, by incubating a dsRNA in
 CC the presence of the proteins. The prepared RNAi agents can be used as
 CC sequence-specific interfering agents for targeted genetic interferences.
 CC The nucleic acids are useful for studying the regulation of RNAi pathway
 CC and to generate knockout strains of animals such as C.elegans. RDE-1 and
 CC RDE-4 genes and their products are useful for modulating RNAi pathway
 CC activity. The polypeptides are useful for generating and testing
 CC antibodies specific for the polypeptides which are useful for studying
 CC the RNAi pathway in C.elegans and other organisms. RNAi pathway genes are
 CC useful for mediating specific processes, e.g. a gene that mediates dsRNA
 CC uptake by the cells is useful for transporting other RNAs into cells or
 CC for facilitating entry of agents such as drugs into cells. RNAi pathway
 CC mutant strains (rde-1, rde-4 mutants) are useful in genetic screens to
 CC identify additional RNAi pathway components
 XX SQ Sequence 3227 BP; 1145 A; 575 C; 666 G; 841 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3227
 Score: 5349.00 Matches: 1020
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-645-746-3 (1-1020) x AAS03283 (1-3227)
 QY 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
 DB 21 ATGTCTCGAATTTCCGAATTTGGAAAGGATTTATCGTCATCTCTCGATCGGAG 80
 QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysVal 40
 DB 81 ATGAAATGGCTTGGAGGCCACTGGTAAATCGCGCCAAATTTCTATGAGAAGAAAGTA 140
 QY 41 LeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTrpGlu 60
 DB 141 CTTCCTTTGGTAAATTTGGTTCAAGTTCTCCAGCAAAATTTACGATCGGGAATACTACCGAG 200
 QY 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
 DB 201 TATGAAGTGAATATGCAAGGAGATTTGATATAGAAACCAAGGAAACCTTTCCCAAAA 260
 QY 81 LysThrGluLeuProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGlu 100
 DB 261 AAGACAGAAATTTCCAATTTCCGATCGTCAAAATCTTTCTGGCAACATCTTCGGCATGAG 320
 QY 101 LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
 DB 321 AAGAAGCAGACAGATTTTATTTCTCGAAGACTATATGTTTTTGATGAAAGGACACTGTGTTAT 380
 QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValLys 140

DB 381 AGTGTCTCGACTGACACTGTACATCAAAAATGCTGCTTTCGGAGAAAGTAGTAAAA 440
 QY 141 LysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTyrThrMetIle 160
 DB 441 AAGGATTCGGAGAAAAAGATTTGGAAGGATTTGGAGAAAAAATCTTATACACAATGATA 500
 QY 161 LeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAspGlu 180
 DB 501 CTTACTTATCGTAAAAAATTTTCCCTGAACTTTAGTCGAGAAAAATCCGGAAGAACGAA 560
 QY 181 GluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysValArgTyrAla 200
 DB 561 GNAAGCATCGAGTTTACAAATTTCTGAAGATGTTATGACCCAGAAAGTTCGTACGCG 620
 QY 201 PropheValAsnGluLysValGlnPheAlaLysAsnPheValTyrAspAsnAsn 220
 DB 621 CTTTTTGTGAACGAGGAGATTTAAAGTACAAATTCGCGAAAAATTTTGTGTACGATAATA 680
 QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeuGlu 240
 DB 681 TCAATTCGCGAGTTCTGGAATCGTTTCCAGATCCAAACAGATTCGAACATCAATTAGAA 740
 QY 241 ValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGluLeuPheAsp 260
 DB 741 GTAGCACCAGAAATCGAAGCATGGTTTGGAAATTTACATTTGAATCAAGAATTTGTCAT 800
 QY 261 GlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMet 280
 DB 801 GGTGAACCTGTCTCAATTTTGCATTTGCGAATAAACTATTCTACAATGCACCGAAAAATG 860
 QY 281 SerLeuLeuAspTyrLeuLeuLeuValAspProGlnSerCysAsnAspAspValArg 300
 DB 861 TCTCTTCTGGATTTATCTTCTCTAAATGTGACCCCGAGTGTGTAACGATGATGTACGA 920
 QY 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro 320
 DB 921 AAAGATCTTAAACCAAACTGATGGCGGAAAAATGACAAATCAGACAAAGCCGCGGCCA 980
 QY 321 ArgIleArgGlnLeuLeuGluAsnLysLysLysCysAlaGluValTrpAspAsnGlu 340
 DB 981 AGAATTCGACAAATTTATTCGAAAAATTTGAAGCTGAAATCGCAGAAAGTTTGGGATACGAA 1040
 QY 341 MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuGluGluAsnSerLeu 360
 DB 1041 ATGTCGATTTGACAGAACGACATCTGACATTTCTAGATTTGTGCGAGGAAACTCTCTT 1100
 QY 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThr 380
 DB 1101 GTTTATAAAGTCACTGGTAAATCGCAGAGAGGAGAAATGCAAAAAAGTACGATACATCA 1160
 QY 381 LeuPheLysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuVal 400
 DB 1161 TTGTTCAAAATCTATGAGGNAACAAAAAAGTTTCATTGAGTTTCCCCCCTACCATAGTC 1220
 QY 401 LysValLysSerGlyValLysGluTyrAlaValProMetGluHisLeuGluValHisGlu 420
 DB 1221 AAAGTTAAAGTGGAGCAAAAGATACCTGTACCAATGGAAACATCTTGAAGTTCAATGAG 1280
 QY 421 LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArg 440
 DB 1281 AAGCCACCAAGATACAAAGATCGAATTCATCTGGTGTATGCAAGCAAGTTCTTAAAGCGA 1340
 QY 441 AlaThrArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAsp 460
 DB 1341 GCTCACGAAAACTTCACGACCTACAAAGAAATACCTTAAAAAATGCTGAAAGAAATGGAT 1400
 QY 461 PheSerSerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMet 480
 DB 1401 TTCCTCTTGAAGAGCTAAATTTTGTGAAAGATTTGGATTTATGCTCCAACTTCAGATG 1460
 QY 481 IleGluCysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGln 500

Db 1461 ATCGAATGTCAGGAAAGGTTTTGAAGAGCCAAATGCTGTGTAATAGTAAATGAACAA 1520
 Qy 501 IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAenValValProGlu 520
 Db 1521 ATTAATAATGACACCAAGTGAATTCGTGGATTTCAAGAAAAACAATGAATGATGCTCCCGAA 1580
 Qy 521 LysGluLeuCysCysAlaValPheValValAenGluThrAlaGlyAenProCysLeuGlu 540
 Db 1581 AAAGAACTTGTCTGCTGTTTTTGTAGTCAACGAAACAGCGGAAATCCATGCTTAGAA 1640
 Qy 541 GluAenAspValValLysPheThrThrGluLeuIleGlyGlyCysLysPheArgGlyIle 560
 Db 1641 GAGAAACAGCTGTGTTAAAGTTCTACACCGAACTAATTCGTGGTCCAGTTCCGTGGAATA 1700
 Qy 561 ArgIleGlyAlaAenGluAenArgGlyAlaGlnSerIleMetThrAspAlaThrLysAen 580
 Db 1701 CGAATTGCTGCCAATGAATAAGAGGAGCGCAATCTATTATGTACGACGCGACGAAAAAT 1760
 Qy 581 GluThrAlaPheThrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAla 600
 Db 1761 GAATATGCCCTTACAAAAAATGTGACACTAAATACCGGAATCGGTAGATTGGAATAGCC 1820
 Qy 601 AlaThrGluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMet 620
 Db 1821 GCAACAGAGCGAAGATATGTTGAACGCTCTCCCGATTAAGAACAAAAAGCTTAATG 1880
 Qy 621 PheIleIleSerLysArgGlnLeuAenAlaThrGlyPheValLysHisThrCysAsp 640
 Db 1881 TTCAATTATCATTTCCAAACGACCACTGAATGCTTACGTTTTGTGAAACATTATTGCGAT 1940
 Qy 641 HisThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAla 660
 Db 1941 CACACCATCGGTAGCTAATACAGCATATTACTTCTGAAACAGTCAAAAGCTTTGGCA 2000
 Qy 661 SerLeuArgHisGluLysGlySerLysArgIlePheThrGlnIleAlaLeuLysIleAen 680
 Db 2001 TCACCTAAGGACGAGAAAGGATCAAAACGAAATTTCTATCAATTCGATTGAAATCAAC 2060
 Qy 681 AlaLysLeuGlyGlyIleAenGlnGluLeuAspThrSerGluIleAlaGluIleSerPro 700
 Db 2061 GCGAAATTAGAGGTATTAACACGAGCTTGACTGGTTCAGAAATTCAGAAATATCACCA 2120
 Qy 701 GluGluLysGluArgArgLysThrMetProLeuThrMetThrValGlyIleAspValThr 720
 Db 2121 GAAGAAAAAGAAAGACGAAACAAATGCCAATTAACCTATGATGTGTTGAATGTGTAAC 2180
 Qy 721 HisProThrSerThrSerGlyIleAspThrSerIleAlaAlaValValAlaSerIleAen 740
 Db 2181 CATCCAACTCTCTACAGTGGAAATGATTTATATAGCGGCTGTAGTAGCGATTCAT 2240
 Qy 741 ProGlyGlyThrIleThrArgAenMetIleValThrGlnGluGluCysArgProGlyGlu 760
 Db 2241 CCAGGTGGAACTATCTATCGAAATATGATGTGACTCAAGAAAGATGTCGCCGGTGAG 2300
 Qy 761 ArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeu 780
 Db 2301 CGTGCACTGCTCATGACGCGGAAGAACAGATATTTTGAAGCAAGTTTCGTGGAATTG 2360
 Qy 781 LeuArgGluPheAlaGluAenAenAspAsnArgAlaProAlaHisIleValValThrArg 800
 Db 2361 CTCAGAGAAATTCGACAGAAACACGACAACTCGAGCACCGCATATTGTAGTCTATCGA 2420
 Qy 801 AspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLys 820
 Db 2421 GACGGAGTTAGCGAATTCGAGATGCTACGTTGTAGTCATGATGAGCTTCGATCTTTAAA 2480
 Qy 821 SerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLysThrThr 840
 Db 2481 AGCGAAGTAAAAACAATTCATGTGCGAACCGGATGGAGAGATCCAGGCGGAGTAGTACCG 2540
 Qy 841 PheIleValIleGlnLysArgHisAenThrArgLeuLeuArgArgMetGluLysAspLys 860
 Db 2541 TTCATTGTGATTCAGAAAGACACATATACGATTGCTTCAAGAAATGGAAAGATATAG 2600

Qy 861 ProValValAenLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys 880
 Db 2601 CCAGTGTGCTCAATAAAGATCTTACTCTCTGCTGAAACAGATGCTGCTGTGCTGCTTAAA 2660
 Qy 881 GlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleValAenProSerSerGly 900
 Db 2661 CAATGGGAGGAGGATATGNAAGAAAGCAAACTGGAATTTGNAACCCATCATCCGA 2720
 Qy 901 ThrThrValAspLysLeuIleValSerLysThrLysPheAspPhePheLeuAlaSerHis 920
 Db 2721 ACACTGTGATTAACCTTATCGTTTCGAAATACAAATTCGAATTTTCTTGGCATCTCAT 2780
 Qy 921 HisGlyValLeuGluThrSerArgProGlyHisThrThrValMetThrAspAspLysGly 940
 Db 2781 CATGGTGTCTCTGTACATCTCGTCCAGGACATTTACACTGTTATGTATGACGATAAGGA 2840
 Qy 941 MetSerGlnAspGluValThrLysMetThrThrGlyLeuAlaPheLeuSerAlaArgCys 960
 Db 2841 ATGAGCCAAAGATGAAGTCTATAAATGACCTCGGACTTGTCTTCTCTCTGTAGATGT 2900
 Qy 961 ArgLysProIleSerLeuProValProValHisThrAlaHisLeuSerCysGluLysAla 980
 Db 2901 CGAAACCCATCTCGTTCCTGCTTCCGTTTCATTTATGCTCATTTATCATGTGAAAAGCG 2960
 Qy 981 LysGluLeuThrArgThrThrLysGluHisThrIleGlyAspThrAlaGlnProArgThr 1000
 Db 2961 AAAGAGCTTTATCGAACTTACAAGGAACATTTACACTGCTGACTATGCACAGCCAGGACT 3020
 Qy 1001 ArgHisGluMetGluHisPheLeuGlnThrAsnValLysThrProGlyMetSerPheAla 1020
 Db 3021 CGACACGAATGAACATTTTCTCCAAACTAACGTGAAGTACCTGGAATGTCTGTCGCA 3080

RESULT 2
 AAS03282
 ID AAS03282 standard; DNA; 3709 BP.
 XX
 AC AAS03282;
 DT 07-SBP-2001 (first entry)
 XX
 DE C. elegans genomic DNA encoding RNA interference pathway protein RDE-1.
 XX RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;
 KW double-stranded RNA-dependent gene silencing; ds.
 XX Caenorhabditis elegans.
 XX Key Location/Qualifiers
 FH CDS 21..3582
 FT /*tag= a
 FT /product= "RDE-1"
 FT /note= "This sequence contains introns"
 XX
 PN WO200129058-A1.
 XX
 PD 26-APR-2001.
 XX
 PF 13-OCT-2000; 2000WO-US028470.
 XX
 PR 15-OCT-1999; 99US-0159776P.
 PR 30-MAR-2000; 2000US-0193218P.
 XX
 PA (UYMA-) UNIV MASSACHUSETTS.
 PA (CARN-) CARNegie INST WASHINGTON.
 XX
 PI Mello CC, Fire A, Tabara H, Grishok A;
 XX WPI; 2001-316239/33.
 DR P-PSDB; AAU01856.
 XX
 PT Novel RNA interference pathway genes and their protein products involved
 in mediation of genetic interference, useful for modulating and studying

PT regulation of RNA interference pathway.

XX Claim 1; Fig 5; 76pp; English.

XX The sequence is genomic DNA encoding the RNA interference (RNAi) pathway protein RDE-1. RDE-1 and RDE-4 are involved in the pathway mediating double-stranded RNA-dependent gene silencing (genetic interference). RDE-1 and RDE-4 protein is useful for preparing an RNAi agent, by incubating a dsRNA in the presence of the proteins. The prepared RNAi agents can be used as sequence-specific interfering agents for targeted genetic interference. The nucleic acids are useful for studying the regulation of RNAi pathway and to generate knockout strains of animals such as C. elegans. RDE-1 and RDE-4 genes and their products are useful for modulating RNAi pathway activity. The polypeptides are useful for generating and testing antibodies specific for the polypeptides which are useful for studying the RNAi pathway in C. elegans and other organisms. RNAi pathway genes are useful for mediating specific processes, e.g. a gene that mediates dsRNA uptake by the cells is useful for transporting other RNAs into cells or for facilitating entry of agents such as drugs into cells. RNAi pathway mutant strains (rde-1, rde-4 mutants) are useful in genetic screens to identify additional RNAi pathway components

XX Sequence 3709 BP; 1308 A; 621 C; 741 G; 1039 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3709
Score: 4804.50 Matches: 1004
Percent Similarity: 84.4% Conservative: 3
Best Local Similarity: 84.2% Mismatches: 13
Query Match: 89.8% Indels: 182
DB: 4 Gaps: 10

US-10-645-746-3 (1-1020) x AAS03282 (1-3709)

QY 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspPro--- 19
DB 21 ATGTCCTCGAATTTTCCCGAATTGGAAAAAGGATTTTATCGTCATCTCTCGATCCGGTA 80
QY 20 -----GluMetIly 22
DB 81 TGAATCAATTTATAGCAGCTATTAAGATATATTAAGTTTGTATTAATATTATAGGAGATGAA 140
QY 22 sTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysLysValLeuLe 42
DB 141 ATGGCTTGGAGCCACCTGGTAAATCCGACGGCAAAATCTATGAGAAAGATGACTTCT 200
QY 42 uLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrG1 62
DB 201 TTTGGTAAATTTGGTTCAAGTTCTCCAGCAAAATTTTACGATCGGGAATACTACGAGTATGA 260
QY 62 uValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysTh 82
DB 261 AGTGAATAATGACAAAAGGAAGTATTGAATAGAAAACCCAGAAAACCTTTCCCAAAAAGA- 319
QY 82 rGluIleProfile-----ProAs 88
DB 320 -GAAATTCANATGAAGTCCTGTAAATTAAGTCAAAACTAATTTTATTTTTCAGTCCCGA 378
QY 88 pArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLysGlnThrAspPheIleLe 108
DB 379 TCGTGCAAAACCTCTTCTGGCAACATCTTCGGCATCAGAAGAAGCAGACAGATTTTATTCT 438
QY 108 uGluAspTyrValPheAspGluLysAspThrValTyrSerValCysArgLeuAsnThrVa 128
DB 439 CGAAGACATATGTTTTTGTATGAAAAGGACATCTGTTATATAGTGTTCGACATGAACACTGT 498
QY 128 lThrSerLysMetLeuValSerGluLysValLysLysAspSerGluLysLysAspG1 148
DB 499 CACATCAAAAATGCTGGTTTCGGAGAAAGTAGTAATAAAGGATTCGGAGAAAAGATGA 558
QY 148 uLysAspLeuGluLysLysIleLeuTyrThrMetIleLeuThrTyrArgLysPheHi 168
DB 559 AAGGATTTGGAGAAAAAAATCTTATACAAATGATACTTACCTATCTCGTAAAAAATTTCA 618

QY 168 sLeuAsnPheSerArgGluAsnProGluLysAspGluGluAlaAsnArgSerTyrLysPhe 188
DB 619 CCTGAATTTTGTAGTCAGAAAAATCCGGAAAAAGACGAAAGCAATTCGGAGTTTACAAAT 678
QY 188 eLeuLys-----AsnValMetTh 194
DB 679 CTTGAAGGTTTATGAAAAACACGCAATTATACAAACAATTAGCTTTTCAGAAATTTATGAC 738
QY 194 rGlnLysValArgTyrAlaProPheValAsnGluGluLysVal----- 209
DB 739 CCAGAAAGTTCTGCTACGCCCTTTTGTGAACGAGGAGATTAAGTGTGAGTTTGAATAAT 798
QY 210 -----GlnPheAlaLysAsnPheVa 216
DB 799 AATAATAATAATCACCTCAACTCATTTATATATATTTTAAGACAAATTCGCGAAAAATTTGT 858
QY 216 lTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheG1 236
DB 859 GTACGATTAATAATTCATTTCTCGAGTTCTCGAATCGTTTCACGATCCAAACAGATTCGA 918
QY 236 uGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLy 256
DB 919 ACAATCATTAGAAGTAGCACCAAGAATCGAAGCATGTTTGAATTTTACATTTGGAATCAA 978
QY 256 sGluLeuPheAspGlyGluProValLeuAsnPheAlaIle----- 269
DB 979 AGAATTTGTTTCGATGTTGAACCTGTCTCAATTTTTCGAAG-TAAGTTTGGAAACTCGCAT 1037
QY 270 -----ValAspLysLeuPheTyrAsnAlaProLysMe 280
DB 1038 AAAAAATCATGTGATTTTGTGAGTTGTCGATAAACTATTCTACATGACCCGGAAT 1097
QY 280 tSerLeuLeuAspTyrLeuLeuLeuLeuValAspProGlnSerCysAsnAspAspValAr 300
DB 1098 GTCTCTTCTGGATATCTTCTCTAAATTTGTGACCCCGACCTGCTGTAACGATGATGACG 1157
QY 300 GlyAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPr 320
DB 1158 AAAAGATCTTAAACAAAAACTGATGCGGGAAAAATGCAATCAGCAAGCCGCGCGGCC 1217
QY 320 oArgIleArgGlnLeuLeuGluAsnLysLeuLysCysAlaGluValTrpAspAsnG1 340
DB 1218 AAGAATTCGACAAATTTTGGAAAAATTTGAAGCTGAATTCGCGCAGAAAGTTTGGGATAACA 1277
QY 340 u-----MetSerArgLeuTh 345
DB 1278 ATGTTAGTTTAAATTTATTTCAAAACAATTAATATACAAATTTGATTTTTCAGGTCGAGATTGAC 1337
QY 345 rGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValTh 365
DB 1338 AGAACGATCTGACATTTCTAGATTTGTGCGAGGAAAACTCTCTGTTTATTAAGTCAAC 1397
QY 365 rGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTy 385
DB 1398 TGGTAAATTCGACAGACAGAGAGAAATGCAAAAAAGTACGATACATACTACATTTTCAAAATCTA 1457
QY 395 rGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerG1 405
DB 1458 TGAGGAAAAACAAAAGTTTCATTGAGTTTCCCACTACCTAGTCAAGTTTAAAGTGG 1517
QY 405 YAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTy 425
DB 1518 AGCAAAAGATACGCTGTACCAATGGAACATCTTGAGTTCTATGAGAGCCCAAAAGATA 1577
QY 425 rLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysPr 445
DB 1578 CAAGAAATCGAATTGATCG-GTGAATGCAAGCAAGTTTCTAAAGCGAGCTTACACGAAAAACC 1636
QY 445 oHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluCl 465
DB 1637 TCAGACTACAAAGAAAAATCCCTAAAAAATGCTGAAAAA-TTGGATTTCTCTTCTGAAGA 1695

465 uLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlu 485
1696 GCTAAATTTTGTGAAAGATTGGATTTATGCTCCAAACTTCAGATGATGAATGTCAGG 1755
485 YLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPr 505
1756 AAAGTTTGAAGAGCAATGCTTGTGATAGTAAATGAACAAATTAATGACACC 1815
505 ovalIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGluLeuCysCy 525
1816 AGTGATTTCGGGATTTCAAGAAACAAATTTGAATGTGTTCCCGAAAGAACTTTCGTC 1875
525 sAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAspValVa 545
1876 TCGTGTGTTTGTAGTCAACGAAACAGCGGAAATCCATGCTTGAAGAGAACGACGTTGT 1935
545 l-----LysPheTyrThrGluLeuI 552
1936 GTAAGTGTGTTTCTACGTAGATTATTCGAAATATTTTCAGTAAGTTCTACCGAACTAA 1995
552 lGlyGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgIleValadInS 572
1996 TTGGTGGTTTCCAAAGTTCCGGTGAATACGAATTTGGTCCAAATGAAACACAGAGGAGCGCAAT 2055
572 erIleMetTyrAspAlaThrLysAsnGlu----- 581
2056 CTATTATGTACGACGCGCAAGAAATGAATATGCGGTAACTTTTCAGAAATTTGAAAGTTT 2115
582 -----TyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgP 597
2116 TTAATATCATATTTACAG-TTCTACAAAATTTGTACACTAAATACCGAATCGGTAGAT 2174
597 heGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnL 617
2175 TTGAATAGCGCAACAGAGCGCAAGAAATATGTTTGAACGCTTCCCGATAAAGAACAAA 2234
617 ysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysH 637
2235 AGTC-TTAATGTTTATTTATTCATTTCCAAACGACACTGAATGCTTACGGTTTTTGTGAAC 2293
637 isTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrL 657
2294 ATTATTCGATCACACCATCGGTAGCTAATCAGCATATTACTTCTGAACAGTACAA 2353
657 ysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaL 677
2354 AAGCTTTGGCATCACTAAGCGACGAGAAAGGATCAAAACGAAATTTCTATCAAAATTCAT 2413
677 eulysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTyrSerGluIleAlaG 697
2414 TGAATAATCAACGCGAAATTAGGAGGTATTAAACAGGAGCTTGACTGGTTCAGAAATTCAG 2473
697 luIleSerProGluGluLysGlu-ArgArgLysThrMetProLeuThrMetTyrValGly 716
2474 AAATATCACAGAGAAAGAAAGAAAGCGAAACAAATGCCATTAATGATGATGTTGA 2533
717 IleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValVal 736
2534 ATTGATGTAATCATCCAACTCC-TACAGTGAATTTGATTTATTTCTATAGCGGTGTAGTA 2592
737 AlaSerIleAsnProGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCys 756
2593 GCGAGTATCAATCCAGTGGAACTATCTATCGAAATATGATGTTGACTCAAGAAGAATGT 2652
757 ArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLys 776
2653 CGTCCCGGTGAGGTGCGTGGCTCATGACGCGAAAGAACAGATATTTTGGAGC-AAA 2711
777 PheValLysLeuLeuArgGluPheAlaGlu----- 786
2712 TTTCGTGAAATTTGCTCAGAGAAATTCGACAGAGTGTGAGTTGCTTCTGAGTATTAAAGATCTC 2771
787 -----AsnAsnAspAsnArgAlaProAlaHisIl 796

2772 TGGGATTTTAAATTTTGTAAACTTTTCAGAACACGACAATCGACCGCATAT 2831
796 eValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLe 816
2832 TGTAGTCTATCGAGCGGAGTTAGCGATTTCGGAGATGCTACGTGTTAGTCATGATGACT 2891
816 uArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGl 836
2892 TCGATCTTTAAAGCGAGTAAACAAATTCATGTCGGAACGGATCGAGAAGATCCAGA 2951
836 uProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgArgMe 856
2952 GCCGAAGTACACGTTCAATGTTGATTTCAGAAAACACACAATACAGATTGCTTCGAAGAAT 3011
856 tGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAlaVa 876
3012 GGAAGAAAGATAAGCCAGT-GTCAATAAAGATCTTACTCTCTGCTGGAACAGATGTCGTGT 3070
876 lAlaAlaValLysGlnTrpGluAspMetLysGluSerLysGluThrGlyIleValAs 896
3071 TGCTGCTGTTTAAACAATGGGAGGAGGATATGAAGAAAGCAAGAAACTGGAATTTGTA 3130
896 nProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhePh 916
3131 CCCATCATCCGGAACCACTGTGATAAACTTATCGTTTCGAAATACAAATTCGATTTTTT 3190
916 eLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTy 936
3191 CTTCGATCTCATCATGTTGTCCTTGGTACATCTCGTCCAGGACATTTACACTGTTATGTA 3250
936 rAspAspLysGlyMetSerGlnAspGluValTyr----- 947
3251 TGACGATAAAGGAATGAGCCAGATGAAGTCTATGTAAGCGTTTTTGAATAGCAGTTAGCG 3310
948 -----LysMetThrT 951
3311 ATTTTAGGATTTTGTAAATCCGCATATAGTTATTATATAAAAAATGTTTCAGAAAATGACCT 3370
951 YrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValH 971
3371 ACGGACTTCTTTCTCTCTGTAGATGTCGAAAACCCATCTCGTTCGCTGTTCCGGTTC 3430
971 isTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHist 991
3431 ATTATGCTCATTTATCATGTGAAAAGCGAAAGAGCTTTATCGAACTTCAAGAACATT 3490
991 YrIleGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThra 1011
3491 ACATCGGTGACTATGCACAGCCACGAGCTCGACACGAAAATGGAACATTTTCTCCAACTA 3550
1011 snValLysTyrProGlyMetSerPheAla 1020
3551 ACGTGAAGTACCTCGAATGTCGTTCCGA 3579

RESULT 3

ADQ80698

ID ADQ80698 standard; DNA; 2960 BP.

XX AC ADQ80698;

XX AC ADQ80698;

XX 21-OCT-2004 (first entry)

XX Arabidopsis thaliana TFL1-binding protein coding sequence #7.

XX Arabidopsis thaliana.

XX TFL1-binding protein; plant growth control; biotechnology;

XX fishing industry; screening; gene; ds.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

XX 1..2958

XX /tag= a

XX FT CDS

XX FT

1554	DB	AATCCAGGCTAGGACAGTGAACATGATTGACAAAGAAATGGTTAATGGAGCAAAAGTC	1713
521	QY	LysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGlu	540
1714	DB	ACTCTTGGACTTGGTAAGTTTCTCTACTCGGATTGACCGTGGTTTACCC	1764
541	QY	GluAsnAspValValLysPheThrThrGluLeuLeuGlyGlyCysLysPheArgGlyLe	560
1765	DB	-----CAAGAGTTCTCGAAACAGTGTGATGGATGTGTGTGCAGCAAGGAATG	1812
561	QY	ArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsn	580
1813	DB	-----GAATTTAGCTCAACCTGCTATTCGGTTCATCTCTGTGTCCTCCCT	1857
581	QY	GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAla	600
1858	DB	GAACAT-----ATTGAGGAAGCT	1875
601	QY	AlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMet	620
1876	DB	CTTCTCGAT-----ATCCAAAGAGGGCACCTGGTCTCCAA-----CTG	1914
621	QY	PheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAsp	640
1915	DB	TTGATTGTAAATTATTCCTCGATGTGCATGGAGTCATATGGAATAAATAAAGATCTGTGA	1974
641	QY	HisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAla	660
1975	DB	ACAGAAATGGGATGTCTCTCAGTGTGCCAACCTAGACAGTTAATAAATC-----	2028
661	QY	SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsn	680
2029	DB	-----AACACGACGATACATGGAAAAATGTTGCTTGAAGATCAAT	2067
681	QY	AlaLysLeuGlyGlyIleAsnGlnLeuAspTrpSerGluIleAlaGluIleSerPro	700
2068	DB	GTCAAGACTGGGGAGGAACACTGTTCTTAAT-----	2100
701	QY	GluGluLysGluArgLysThrMetProLeu-----ThrMetTyrVal	715
2101	DB	-----GATGCTATTAGAGAAACATACCTTTATTACTGATCGTCCAAACATCATCATG	2154
716	QY	GlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaVal	735
2155	DB	GGTCTGATGTGACTCACCCACAGCTGGAGAGGACTCAAGTCTCTTATTGCTGCTGTT	2214
736	QY	ValAlaSerIleAsn-----ProGlyGlyThrIleTyrArgAsnMetIleValThrGln	753
2215	DB	GTGGCTCTATGGACTGGCTTGAGATAAACAAATACCGAGGNTGGTTTTCTGCTCAAGCT	2274
754	QY	-----GluGlu-----CysArgProGlyGluArgAlaVal	763
2275	DB	CATAGGGAAGAAATTATTTCAGGACCTGTATAGCTGGTTCAGGATCCAAACGTGGGCTA	2334
764	QY	AlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGlu	783
2335	DB	GTCCAC-----TCTGGTTTGTAAAGGGAACATTTTCATGATTCAGGAGAGCT	2382
784	QY	PheAlaGluAsnAsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyVal	803
2383	DB	-----ACAGCCACAGATACCTCAAGGATCATCTTCTATCGTGACGGAGTA	2427
804	QY	SerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluVal	823
2428	DB	AGCGAGGGCAGTTTACTCAGTTTCTGCTACATGAGATGACTGCTATATCCCAAGGCTTGT	2487
824	QY	LysGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleVal	843
2488	DB	AACTCTCTCCAA-----GAGAATTATGTTCTCTCGTGTACTTTTCGTGATT	2532
844	QY	IleGlnLysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValVal	863
2533	DB	GTCCAGAAAGCTCACCAACACCTGTTGTTCCCTGAGCAACACGGG-----	2577

Qy	864	AsnLysAspLeuThrProAlaGluThrAspValaValaValaValaValysGlnTrpGlu	883
Db	2578	AATCGTGATATGACT	2599
Qy	884	GluAspMetLysGluSerLysGluThrGlyLeValaAsnProSerSerGlyThrThrVal	903
Db	2593	-----GATTAAGATGGCAATATTCAACCA-----GGTACTGTCTG	2628
Qy	904	AspLysLeuLeuValSerLysPheAspPhePheLeuAlaSerHisGlyVal	923
Db	2629	GACACTAAATCTGTCAACCCTAATGAATTCGACTTCTATTATGAACGCATGCTGGTATT	2688
Qy	924	LeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGln	943
Db	2689	CAGGGAACACAGCAGCGCGCACATTACCATGTACTTCTCGATGAGAACGGTTCACCGCT	2748
Qy	944	AspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysPro	963
Db	2749	GATCAGTTGCCAATGCTCACAAACACCTCTGCTACAGTATCGGAGGTGTACAAAATCT	2808
Qy	964	IlSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeu	983
Db	2809	GTGTCAATTGGCCACCAGCGCTACTACGCTCACTTGGCTGCATTCCGTGCC-----	2859
Qy	984	TyrArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThrArg	1001
Db	2860	---CGCTACTACATGAGAGTGAGATGCTCGATGGAGGTTTCGAGCAGGTCCAGG	2910
RESULT 4			
ID	ADJ94703	standard; cDNA; 2571 BP.	
AC	ADJ94703;		
DT	06-MAY-2004	(first entry)	
XX	Human eukaryotic translation initiation factor 2C1 (eIF2C1) cDNA.		
KW	cytostatic; antiinflammatory; virucide; immunosuppressive; tumour;		
KW	inflammatory; infectious disease; viral infection; degenerative;		
KW	autoimmune; gene therapy; Argonaute family;		
KW	eukaryotic translation initiation factor 2C1; eIF2C1; human; ss; gene.		
OS	Homo sapiens.		
XX	WO2004007718-A2.		
XX	22-JAN-2004.		
XX	10-JUL-2003; 2003WO-EP007516.		
XX	10-JUL-2002; 2002EP-00015532.		
PR	23-AUG-2002; 2002EP-00018906.		
XX	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.		
PA	Tuschl T, Martinez J, Patkaniowska A, Urlaub H, Luehrmann R;		
XX	WPI; 2004-122948/12.		
DR	P-PSDB; ADJ94697.		
XX	New single-stranded RNA molecule having a length from 14-50 nucleotides,		
PT	useful for preventing or treating tumor, inflammatory, infectious, e.g.		
PT	viral infections, degenerative and autoimmune diseases.		
XX	Example; Fig 16; 73pp; English.		
XX	The invention relates to a novel single-stranded RNA molecule having a		
CC	length from 14-50 nucleotides where at least 14-20 of the 5'-most		
CC	nucleotides are substantially complementary to a target transcript. The		
CC	RNA molecule of the invention demonstrates cytosstatic, antiinflammatory,		
CC	virucide and immunosuppressive activities and may be useful for		

CC inhibiting the expression of a target gene in vitro or in vivo, the
 CC preferably for preventing or treating diseases associated with the
 CC overexpression of at least one target transcript. The diseases may be
 CC selected from tumour diseases, inflammatory diseases, infectious diseases
 CC such as viral infections, degenerative diseases and autoimmune diseases.
 CC Furthermore, the molecules of the invention may be utilised during gene
 CC therapy. The current sequence is that of the human eukaryotic translation
 CC initiation factor 2C1 (eIF2C1) cDNA of the invention.

XX
 SQ Sequence 2571 BP; 628 A; 717 C; 690 G; 536 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-36e-43 Length: 2571
 Score: 577.00 Matches: 231
 Percent Similarity: 40.8% Conservatives: 171
 Best Local Similarity: 23.4% Mismatches: 361
 Query Match: 10.8% Indels: 224
 DB: 12 Gaps: 37

US-10-645-746-3 (1-1020) x ADJ94703 (1-2571)

QY 38 LysLysValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGlu 57
 DB 100 AAACAATCAAGCTCTCGCCCAATTACTTTGAGGTGGACATCCCTAAGATCGACGTGTAC 159
 QY 58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro 77
 DB 160 CACTACGAGGTGGACATCAAGCCGGATAAG----- 189
 QY 78 PheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
 DB 190 TGTCCCGTAGAGTCAACCGGAGTGGTGGA-----TACATGGTTC 231
 QY 98 ArgHisGluLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGlu 115
 DB 232 CAGCATTTCAAGCCTCAG-----ATCTTTGGTGATCGCAAGCCTGTGTATGGA 282
 QY 116 LysAspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSer 135
 DB 283 AAGAAGAACATTAC-----ACTGTCAACAGCATGCCCATTTGGCAAC 324
 QY 136 GluLysVal-----ValLysLysAspSerGluLysLysAspGluLysAspLeuGlu 152
 DB 325 GAACGGTTCAGCTTTGAGGTGACATCCCTGGGGAAGGAAGAT---CGAATCTTAAG 381
 QY 153 LysLysIleLeuTyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSer 172
 DB 382 GTCTCCATCAAGTGGCTAGCCATTGTGAGCTGGCGAATGCTGCATGAGGCCCTGTGTGAC 441
 QY 173 ArgGluAsnProGluLysAspGluAlaAsnArgSerTyrLysPheLysAsnVal 192
 DB 442 GGCAGATCCCTGTTCCTTGGAGTGTGTGCAAGCCTGGATGTGGCCATGAGGACCTG 501
 QY 193 MetThrGlnLysValArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAla 212
 DB 502 GCATCC-----ATGAGGTACACCTCTGTG----- 525
 QY 213 LysAsnPheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspPro 232
 DB 526 -----GGCGGCTCTTCTTCTCACCGCTGAGGGCTACTACCACCG 567
 QY 233 AsnArgPheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyr 252
 DB 568 -----CTGGGGGTGGCGGAGGTCTGGTCTGGCTTTTAC 603
 QY 253 IleGlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLys 272
 DB 604 CAGTCTGTGGCCCTGCCATG---TGAAGATGATGCTCAACATTGATGTCTCAGCCACT 660
 QY 273 LeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspPro 292
 DB 661 GCCTTTTATAGGCA-----CAGCCAGTGATTGATTCATGTGTGAGGTGTGGACATC 714

QY 293 GlnSerCysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMet 312
 DB 715 AGGAACATAGATGACAGCCCAAG----- 738
 QY 313 ThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys 332
 DB 739 CCCCTCAGGACTCTCAGCGGCTTCGCTTCCACCAAGGAGATCAAGGCCCTGAAGGTGAA 798
 QY 333 -----CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThr 345
 DB 799 CTCACCCACTGTGGACATGAAGAGGAGTACCCGCTGTGTAATGTATACCCCTGCCCT 858
 QY 346 GluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThr 365
 DB 859 GCTAGCCATCAGACATTTCCCTTACAGCTGGAGAGTGGACACTGTG----- 906
 QY 366 GlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyr 385
 DB 907 -----GAGTGCACAGTGGCACAGTATTTC 930
 QY 386 GluGluAsnLysLysPhe-----IleGluPheProHisLeuProLeuValLysValLysSer 404
 DB 931 AAGCAGAAATATAACCTTCAGCTCAAGTATCCCATCTGCCCTGCTTACAGTTGCCAG 990
 QY 405 GlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArg 424
 DB 991 GAACAAAGCATACCTACTCTTCCCTAGAGGTCTGTAACTT---GTGGCTGGGCGCCG 1047
 QY 425 TyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLys 444
 DB 1048 TGTATTAAGAGCTGACCGACAACCCAGACCTCGCATGATAAAGCCACAGCTAGATCC 1107
 QY 445 ProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGlu 464
 DB 1108 GCTCCAGACAGACAGGAGGAGATCAGTCCCTGATGAAGAATGCCAGCTACAACTTAGAT 1167
 QY 465 GluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysPro 484
 DB 1168 CCC-----TACATCCAGGAATTTGGATCAAGTGAAGGATGACATGACGGAGGTGACA 1221
 QY 485 GlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThr 504
 DB 1222 GGGCGAGTGTCTCGCGGCCCATCTTGCAGTACGGCGCGCGCAACCGGCCCATTTGCCACA 1281
 QY 505 Pro-----ValIleArgGlyPheGlnGluLysGlnLeu---AsnValVal 518
 DB 1282 CCCAATCAGGTGTCTGGGACATCGGGGG-----AAACAGTTCTACATGGGATT 1332
 QY 519 ProGluLysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCys 538
 DB 1333 GAGATCAAGTCTGGGCCATCGCTGCTTCGCACCCCAAAA-----CAGTCT 1380
 QY 539 LeuGluGluAsnAspValLysPheTyrThrGluLeuIle-Gly----- 553
 DB 1381 CGAAGAAG-----GTGCTCAAGAACTTTCACAGACCTCGGGAAGATTTCCAAGGAT 1434
 QY 554 -----GlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSe 572
 DB 1435 GCGGGATGCTTATCCAGGGTCAACCTTTGTCGAAATATGCACGGGGGCA-CACAG 1493
 QY 572 rIleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnTh 592
 DB 1494 CGTG----- 1497
 QY 592 rGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuPr 612
 DB 1498 -----GAGCCTATGTTCCGCATCTCAA 1520
 QY 612 AspLysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTy 632
 DB 1521 GAACACCTACTCAGGGCTGCGAGCTCATTTGTTCATCTCTGCCAGGGAAGACCCCGGTGTA 1580
 QY 632 rGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSe 652

US-10-645-746-3 (1-1020) x ADB81486 (1-7478)

QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu----- 37
DB 209 ATGGGATGAAA-GCGGACCCCTGGGAGCAGCTGCGGCGCTTACCTGCCCCCCTGCAG 267
QY 38 -----LysLysValLeuLeu 42
DB 268 CAGGTGTTCCAGGCACCTCGCGGCTGGCANTGGCAGCTGTGGGGAACCAATCAAGCTC 327
QY 43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgLysPheTyrGlu 62
DB 328 CTGGCCAAATTACTTTGAGGTGGACATCCCTAAGATCGACGTGTACCCTACGAGGTGGAC 387
QY 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82
DB 388 ATCAAGCCGGATAAG-----TGTCCTCCCGTAGAGTC 417
QY 83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
DB 418 AACCGGAAGTGGTGGAA-----TACATGGTCAGCAATTTCAAGCCT 459
QY 103 GlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLysAspThrValTyr 120
DB 460 CAG-----ATCTTTGGTGTATCGAAGCCCTGTGTATGATGGAAGAAGAACATTTAC 510
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal----- 138
DB 511 -----ACTGTACAGCAGCTGCCCATTTGGCAACAGCGGTGCACTTT 552
QY 139 ---ValLysLysAspSerGluLysLysAspGluLysAspLeuLysLysLysLysLys 157
DB 553 GAGGTGACCAATCCCTGGGAGAGGAAGGAT---CGAATCTTTAAGTCTCCATCAAGTGG 609
QY 158 ThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
DB 610 CTAGCCATTTGTAGCTGGCGAATCTGTCATGAGCGCTGTGTAGCGCCAGATCCCTGT 669
QY 178 LysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysVal 197
DB 670 CCTTGGAGTCTGTGCAAGCCCTGGATGTGGCATGAGGCACCTGGCATCC-----ATG 723
QY 198 ArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTyr 217
DB 724 AGGTACACCCCTGTG----- 738
QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
DB 739 ---GGCGCTCTCTTCTCAGCGCTGAGGGCTACTACACCCG----- 780
QY 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyLysGlu 257
DB 781 -----CTGGGGGTGGCGCGAAGTCTGTTCTGGCTTTCACCAGTCTGTGGCCCT 831
QY 258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
DB 832 GCCATG---TGGAGATGATGCTCAACATGTATGTCTCAGCCACCTGCTTTTATAGGCA 888
QY 278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAsp 297
DB 889 -----CAGCCAGTATGATTGATGTCATGTGTGAGTGTGTGACATCAGGAACATAGTAG 942
QY 298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
DB 943 CAGCCCAAG-----CCCTCAGGACTCT 966
QY 318 AlaAtProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys-----CysAla 334
DB 967 CAGCGGTTTCGCTTCCACAGGAGATCAAGGGCTTGAAGTGAAGTCAACCCACTGTGGA 1026
QY 335 GluValTrpAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr 350
DB 1027 CAGATGAAGAGGAAGTACCGCGTGTGTAATTTACCCTGCGCCTGCTAGCCATCAGACA 1086

QY 351 PheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArg 370
DB 1087 TTCCTCTTACAGCTGAGAGTGGACAGACTGTG----- 1119
QY 371 GlyArgAsnAlaLysLysTyrThrLeuPheLysIleTyrGluGluAsnLysLys 390
DB 1120 -----GAGTGCACAGTGGCACAGTATTTCAAGCAGAAATATAAC 1158
QY 391 Phe-----IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyr 409
DB 1159 CTTAGCTCAAGTATCCCATCTGCCCTTACAAAGTTGGCCAGGACAAACAAAGCATACC 1218
QY 410 AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIle 429
DB 1219 TACCTTCCCTAGAGGTCTGTAAACAT---GTGGCTGGCAGCGCTGTATTAAAGAGCTG 1275
QY 430 AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys 449
DB 1276 ACCGCAACACAGACCTCGACCATGATAAAGGCCACAGCTAGATCCGCTCCAGACAGACAG 1335
QY 450 GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal 469
DB 1336 GAGGAGATCAGTCGCTGATGAAGATGCCAGCTACAACCTTAGATCCC-----TACATC 1389
QY 470 GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
DB 1390 CAGGAATTTGGATCAAAAGTAGATGACATGACGAGGTGACAGGGCGAGTGTGCGG 1449
QY 490 GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro----- 505
DB 1450 GCGCCCATCTTGCAGTACGCGCGCGGAACCGGGCCATTTGCCACACCCCAATCAGGTGTC 1509
QY 506 ---ValIleArgGlyPheGlnGluLysGlnLeu-----AsnValValProGluLysGluLeu 523
DB 1510 TGGGCATGCGGGGG-----AAACAGATTCTCAATGGATTTGAGATCAAAAGTCTGG 1560
QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
DB 1561 GCCATCGCTGCTTCGACCCCCCAAAA-----CAGTGTGAGAAGAG----- 1602
QY 544 ValValLysPheTyrThrGluLeuIleGly-----GlyCysLysPhe 557
DB 1603 GTGCTCAAGAACTTCACAGACAGCTGCGGAAGATTTCCAAGGATGCGGGGATGCTATC 1662
QY 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
DB 1663 CAGGCTCAACCTTGTTCGAAATATGCACAGGGGGCA-GACAGGCTG----- 1710
QY 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPhe 597
DB 1710 ----- 1710
QY 597 eGluIleAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLys 617
DB 1711 -----GAGCTATGTTCCGGCATCTCAAGAACACCTTACTCAGG 1748
QY 617 sValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis 637
DB 1749 GCTCAGCTCATATTGTTCATCTCTGCCAGGAAGACGCCGGTGTATGCTGAGGTGAACG 1808
QY 637 sTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
DB 1809 TGTCCGAGATACACTCTTGGGAATGGCTACGAGTGTGTGAGGTGAAGAACGTGTGCAA 1868
QY 657 sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLe 677
DB 1869 GACC-----TCACCTCAGACTCTGTCCAACTCTGCTC 1901
QY 677 uLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGln 697
DB 1902 CAAGATCAATGTCAAACTTGGTGGCATTAACAACTCCTA----- 1941

QY 697 uileSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyI 717
 Db 1942 ----GTCCACACAGCGCTCTGCGTTTTCACAGCCAGTG---ATATTCTGGGAGC 1994
 QY 717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAl 737
 Db 1995 AGATGTTTACACACCCCGCAGCGGGATGGAAAAACCTTCTATCACAGCATGGTAGG 2054
 QY 737 aSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGlyCysAr 757
 Db 2055 CAGTATGGATGCCACCCAGCCGATATCTGTGCTACTGTGGGTACAG-----2103
 QY 757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
 Db 2104 -----CGACACCGGCAAGATCATTTGAAGACTGTGTC 2135
 QY 775 aLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794
 Db 2136 CTACATGTTGGTGAGCTCTCTCATCCCAATTTCTACAGTCCACCGTTTCAAG---CTTAC 2192
 QY 794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
 Db 2193 CCGCATCATCTTCTACCGAGATGGGTGCTGAAAGCGCAGCTACCCAGATACCTCCATTA 2252
 QY 814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834
 Db 2253 TGAGCTACTGGCCATCTCGTGATGCTGCATCAAACTG-----GAAAGGA 2297
 QY 834 pProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuAr 854
 Db 2298 CTACAGCCTGGGATCATTTATATTTGGTGCGAAGAGCGCATCACACCGCTTTC-- 2355
 QY 854 gArMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspVa 874
 Db 2356 -TGTCCTGACAGAATGAGCGAATGGGAAGAGTGGTAACATCCAGCT-----2403
 QY 874 lAlaValAlaValLysGlnTrpGluGluAspMetLysLysLysGluThrGlyI 894
 Db 2403 -----2403
 QY 894 eValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAs 914
 Db 2404 -----GGGACACAGTGGACACCAATCACCCACCATTTGAGTTGA 2447
 QY 914 pPheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrVa 934
 Db 2448 CTTCTATCTGTGCAGCCACGAGCATCCAGGCGCACCGCCGACATCCCATTTACTATGT 2507
 QY 934 lMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAl 954
 Db 2508 TCTTTGGGATGACAACCGTTTTCACAGATGAGCTCCAGATCCTGACGTACAGCTGTG 2567
 QY 954 aPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHi 974
 Db 2568 CCACACTTACGTACGATGCACAGCTCTGTCTTATCCAGACCTGCTCTATATGCCCCG 2627
 QY 974 sLeuSerCysGluAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAs 994
 Db 2628 CCTGTGGCTTTCCGGGACGA---TACCACCTGTGGTGGACAGAGCATGACAGTGGAGA 2684
 QY 994 p 994
 Db 2685 G 2685
 RESULT 6
 ID AEC01623
 XX AEC01623 standard; DNA; 7478 BP.
 AC AEC01623;
 XX
 DT 20-OCT-2005 (first entry)
 XX
 DE Human EIF2C1 DNA, SEQ ID NO: 3.

XX Cancer; cytostatic; neoplasm; hypercholesterolemia; antilipemic;
 KW metabolic disorder; gene; ds; EIF protein kinase family;
 KW eukaryotic translation initiation factor protein kinase family;
 KW chromosome 1.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 214..2787
 FT /*tag= a
 FT /product= "Human EIF2C1 protein"
 XX
 PN US2005182015-A1.
 XX 18-AUG-2005.
 PD 14-JAN-2005; 2005US-00035669.
 PF 23-FEB-2001; 2001US-00793807.
 PR 12-SEP-2001; 2001US-00954679.
 PR 13-SEP-2001; 2001US-00953611.
 PR 08-NOV-2001; 2001US-00007078.
 XX (WARD/) WARD D T.
 PA (WATT/) WATT A T.
 XX Ward DT, Watt AT;
 PI WPI; 2005-563220/57.
 DR P-PSDB; AEC01709.
 DR REFSEQ; NM_012199.2.
 XX New antisense oligonucleotides which inhibits expression of eukaryotic
 PT initiation factor 2C1, useful for modulating RNA interference and
 PT treating a disease or condition characterized by hypercholesterolemia, or
 PT cancer.
 XX
 PS Example 13; SEQ ID NO 3; 67pp; English.
 CC The present invention relates to an antisense oligonucleotide which
 CC specifically hybridizes with the polynucleotide encoding eukaryotic
 CC translation initiation factor 2C1 (EIF2C1; also known as Co-EIF-2C,
 CC eIF2C, golgi ER protein 95kDa, GERP95 and Q99) and inhibits its
 CC expression. The invention is useful for treating hypercholesterolemia and
 CC hyperproliferative disorder such as cancer. The present sequence is the
 CC human EIF2C1 DNA which is located on chromosome 1p34-p35.
 XX
 SQ Sequence 7478 BP; 1750 A; 1863 C; 1895 G; 1970 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6,06e-43 Length: 7478
 Score: 577.00 Matches: 237
 Percent Similarity: 40.2% Conservative: 173
 Best Local Similarity: 23.2% Mismatches: 370
 Query Match: 10.8% Indels: 243
 DB: 14 Gaps: 38
 US-10-645-746-3 (1-1020) x AEC01623 (1-7478)

QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu----- 37
 Db 209 ATGGGATGGAA-GCGGGAGCCCTCGGAGCAGCTGCGGCGCTTACCTGCCCCCTCGAG 267
 QY 38 -----LysLysValLeuLeu 42
 Db 268 CAGGTGTTCCAGGCACCTCGCGGCTCGCATTTGGCATTGGGGAACCAATCAAGCTC 327
 QY 43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGlu 62
 Db 328 CTGGCCAATATTCTTGAGGTGGACATCCCTAAGATCGACGTGTACCACTACGAGGTGAC 387
 QY 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82

QY 78 PheProLysThrGluLeuProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
Db 570 ----- 570
QY 98 ArgHisGluLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp 117
Db 570 ----- 570
QY 118 ThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLys 137
Db 571 ----- 609
QY 138 ValValLysLeuAspSerGlu-----LysLysAspGluLysAspLeuLysLys 154
Db 610 ----- 660
QY 155 IleLeuTyrThrMet-----IleLeuThrTyrArgLysLysPheHisLeuAsnPheSer 172
Db 661 AGCCTCTATACCTGCTGGTCCATTACCTTTTGACTCGAAAGAGTGTGTGTGAATCTGGCG 720
QY 173 ArgGluAsnProGluLysAspGluGluAlaAsnArgSerTyrLysPhe-----LeuLysAsn 191
Db 721 GAGAAAAGAGCTGACCGTTCCTCTGGGAAGGACAGACCGTTTAAAGTTGCTGTAAGAAT 780
QY 192 Val-----MetThrGlnLysValArgTyrAla 200
Db 781 GTGACAGACGCTGATCTTTATCAGTTGCAACAGTTCCTTGATCGTAAGCAAGAGGCT 840
QY 201 ProPheValAsnGluGluLysValGlnPheAlaLysAsnPheValTyrAspAsnAsn 220
Db 841 CCATAT-----GATACTATCAAGTGTGTGATGTTGTTTGGGTAAGCCCTCTAAT 894
QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeu--- 239
Db 895 GATTATGCTCTGTTGGGAGGCTTTT-----TTCACACTAGTATTGGGA 939
QY 240 -----GluValAlaProArgIleGluAlaTrpPheGlyIle 251
Db 940 AAGGAGCGAAGAGATGTTAGGGGTGAGCTTGAGATGGTATTGTAGTCTGGAGAGGTTAT 999
QY 252 TyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAsp 271
Db 1000 TTCAAAGTCTAAGG---CTGACTCAGATGGGTTGTCTCTGAACATTGACGTTTCAGCA 1056
QY 272 LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAsp 291
Db 1057 AGATCATTTTAT-----GAA 1071
QY 292 ProGlnSerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
Db 1072 CCGATTGTGTCAGTGACTTTATTATAGCAAGTTTCTGAATATAGGGACTTA-----AAC 1125
QY 312 MetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeu 331
Db 1126 AGGCCACTTAGAGACTCAGACTCGACTTAAGGTGAAGAAAGTTTTCAGGACACTGAAGTT 1185
QY 332 LysCysAlaGluValTrpAspAsnGluMetSerArg----- 343
Db 1186 AAGTTGCTTTCAC---TGGAAACGGCAGCAAAAGATGCGCAAAATTTAGTGGGATTTCTAGTCTA 1242
QY 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyr 362
Db 1243 CCCATCAGGAGCTAAGTTCTCTGGAGGAGCAATCAGAGACGCGTTGTTCAATAT 1302
QY 363 LysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPhe 382
Db 1303 -----TTTGCTGAAAAATAAAT 1320
QY 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
Db 1321 -----TATAGA-----GTGNAATACAGGCTCTACCTGCTATTCAACA 1359
QY 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422

Db 1360 GGGAGTGACACAAGACCCGCTTACCTACCAATGGAGCTCTGCCAAATTGACGAA---GGG 1416
QY 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
Db 1417 CAAGATACACCAAAAGGCTCAATGAGAAGCAAGTGACTGCTGCTAAAGACTACCTGC 1476
QY 443 ArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSer 462
Db 1477 CAACACCCCTGAT---AGAGAGAACCTCGATCAAAAACCTTGTGTTGTGAAAAATATTAC 1533
QY 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu 482
Db 1534 AATGATGATCTGAGC-----AAGGAGTTGGGATGTCAGTGACTACCACTAGCTGCTG 1587
QY 483 CysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLys 502
Db 1588 ATTGAAGCTCGTGACTTCCCCACCGGATGTTGAAGTAGTACCATGACAGTGTGAAGAGAAA 1647
QY 503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGlu 522
Db 1648 ATGTTAAATCCAAGGCTAGGA-----CAGTGAACATGATGACAGAAA--- 1692
QY 523 LeuCysCysAlaValPheValAsnGluThrAlaGlyAsnProCysLeuGluGluAsn 542
Db 1692 ----- 1692
QY 543 AspValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheArgGlyIleArgIle 562
Db 1693 -----ATGTTTAAATGGAGCAAAA---GTCACTTCT 1719
QY 563 GlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyr 582
Db 1720 TGGACTTCGGAATTTAAGCCTCAACCTGCTATTCGGTTCATCTCTTGCCCCCTGAACAT 1779
QY 583 AlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaIleThr 602
Db 1780 -----ATTGAGGAAGCTCTTCTC 1797
QY 603 GluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle 622
Db 1798 GAT-----ATCCAAAAAGGCGCCTGCTCTCCAA---CTGTGTGATT 1836
QY 623 IleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAspHisThr 642
Db 1837 GTATATTGCTGATGCTGCTGATCATATGGAATAATAAAGGATCTGTGAAACAGAA 1896
QY 643 IleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeu 662
Db 1897 TTGGGATGCTCTCTCAGTGTGCCCACTAGACAAAGTTAATAAACTC----- 1944
QY 663 ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLys 682
Db 1945 -----AACAAACGATCATGGAATAATGTCCTTGAAGATCAATGTCAAG 1989
QY 683 LeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGluIleSerProGluGlu 702
Db 1990 ACTGGGGAAGGAACACTGTTCTTAAT----- 2016
QY 703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
Db 2017 GATGCTATTAGAAGAAACATACCTCTTATTACTGATCGTCCAACTCATCATGCGTGTG 2076
QY 718 AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAla 737
Db 2077 GATGTGACTCACCACACCCCTGGAGAGGACTCAAGTCTCTTATTGCTGCTGTGTGGCC 2136
QY 738 SerIleAsn---ProGlyGlyThrIleTyrArgAsnMetIleValThrGln----- 753
Db 2137 TCTATGAGTGGCTGAGATAACAATACCGAGGATGGTTTCTGCTCAAGCTCATAG 2196
QY 754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
Db 765 -----

```

Db 2197 GAAGAAATTATTACGACCTGTATTAAGCTGGTTCCAGGATCCACAACGTGGGCTAGTCAC 2256
Qy 766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
Db 2257 -----TCTGGTTTGTATAAGGGAACATTTTCATAGCATTCAGGAGAGCT----- 2298
Qy 786 GluAsnAsnAspAsnArgAlaProAlaHisIleValValTyArgAspGlyValSerAsp 805
Db 2299 -----ACAGCCAGATACCTCAAGAGGATCATCTTCTATCGTACGGAGTAGCGAA 2349
Qy 806 SerGluMetLeuArgValSerHisAspGluLeuLeuArgSerLeuLysSerGluValLysGln 825
Db 2350 GGGCAGTTAGTCAGGTTCTGCTACATGAGATGCTGCTATCCGACAGGCTTGTAACTCT 2409
Qy 826 PheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGln 845
Db 2410 CTCGAA-----GAGAATTATGTTCTCTCGTTACTTTCGTGATGTGTCAG 2454
Qy 846 LysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLys 865
Db 2455 AAACGTCACACACACAGCTTTGTTCCCTGAGCAACACCGG-----AATCGT 2499
Qy 866 AspLeuThrProAlaGluThrAspValAlaValAlaAlaValLysGlnTrpGluLysAsp 885
Db 2500 GATATGACT----- 2508
Qy 886 MetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLys 905
Db 2509 -----GATAGAGTGGCAATATTCACCA-----GGTACTGCTGGACACT 2550
Qy 906 LeuIleValSerLysTyrLysPheAspPheLeuAlaSerHisIleGlyValLeuGly 925
Db 2551 AAAATCTGTCACCCATAAGTAATTCGACTTCTATTGAAACGCCATGCTGTATTTCAGGGA 2610
Qy 926 ThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGlu 945
Db 2611 ACAAGCAGCGCGCACATTACCATGTTACTTCTCGATGAGACGGTTTCACCGCTGATCAG 2670
Qy 946 ValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSer 965
Db 2671 TTGCAATGCTCACAACAACCTCTGCTACACGTATCGAGGTGTACAAAATCTGTGTCA 2730
Qy 966 LeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArg 985
Db 2731 ATTGTGCCACCGCTACTACGCTCACTTGGCTGCTATTCGTCGCC-----CGC 2778
Qy 986 ThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThrArg 1001
Db 2779 TACTACATGGAGATGAGATGTCGTGATGGAGTTCGAGCAGGTCACGG 2826

RESULT 8
ID ADC46702
XX ADC46702 standard; DNA; 2910 BP.
XX
XX ADC46702;
XX
XX 18-DEC-2003 (first entry)
XX
XX Thalecress transcription factor-like DNA G1149.
XX
XX Thalecress; transcription factor-like protein; ds; seed trait;
XX transgenic; gene; plant size; stress tolerance; yield;
XX disease resistance; plant.
XX
XX Arabidopsis thaliana.
XX
XX US2003093837-A1.
XX
XX 15-MAY-2003.
XX
XX 01-NOV-2002; 2002US-00286264.
XX
XX 23-MAR-1999; 99US-0125814P.
XX

```

22-MAR-2000; 2000US-00533030.

(KEDD/) KEDDIE J.
 (RIEC/) RIECHMANN J L.
 (RATC/) RATCLIFFE O.
 (ZHAN/) ZHANG J.
 (JIANG/) JIANG C.
 (PINE/) PINEDA O.
 (HEAR/) HEARD J.
 (YUGG/) YU G.
 (ADAM/) ADAM L.
 (BROU/) BROUN P.
 (REUB/) REUBER L.
 (PILG/) PILGRIM M.
 (SAMA/) SAMAHA R.

Keddle J, Riechmann JL, Ratcliffe O, Zhang J, Jiang C, Pineda O;
 Heard J, Yu G, Adam L, Broun P, Reuber L, Pilgrim M, Samaha R;

WPI; 2003-765498/72.
 P-PSDB; ADC46703.

Novel transgenic plant having recombinant polynucleotide encoding
 polypeptide that alters trait of transgenic plant when compared with same
 trait of another plant lacking recombinant polynucleotide.

Disclosure; SEQ ID NO 101; 165pp; English.

The invention relates to a transgenic plant having recombinant
 polynucleotide (II) encoding polypeptide comprising at least 6
 consecutive amino acids of a sequence chosen from the protein sequence
 appearing as ADC46603 - ADC46749 (every second sequence), where
 recombinant polynucleotide alters a trait of the seed transgenic plant
 when compared with same trait of another plant lacking recombinant
 polynucleotide. The proteins are transcription factor-like proteins. Also
 included are altering (M1) a trait associated with seed (comprising:
 transforming a plant with (II); selecting the transformed plants; and
 identifying a transformed plant with seed having altered trait); altering
 (M2) the expression levels of at least one gene of a plant (involving
 transforming the plant with (II) and selecting the transformed plant),
 altering (M3) a trait associated with a plant's seed (comprising:
 transforming the plant with a recombinant polynucleotide comprising a
 nucleotide sequence comprising least 18 consecutive nucleotides of a
 sequence appearing as ADC46750 - ADC46766 and selecting the transformed
 plant) altering (M4) a plant's trait (involving providing a database
 sequence, comparing the database sequence with a polypeptide or a
 polynucleotide chosen as detailed above, selecting a database sequence
 that needs selected sequence criteria and transforming a database
 sequence in the plant) and altering a plant's trait (involving providing
 a test polynucleotide, hybridising the test polynucleotide with a
 polynucleotide detailed above at low stringency and transforming the
 hybridising test polynucleotide in a plant to alter a trait of the
 plant). The method (M1) is useful for altering a trait associated with
 seed. The method (M2) is useful for altering the expressing levels of at
 least one gene of a plant. The method (M3) is useful for altering a trait
 associated with a plant's seed. The method (M4) is useful for altering a trait
 plant's trait. The method (M4) is useful for altering a plant's trait
 such as seed or plant size, stress tolerance, yield or disease
 resistance. The present sequence encodes a transcription factor-like
 protein/seed trait altering protein of the invention.

Sequence 2910 BP; 825 A; 590 C; 724 G; 771 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2-24e-43	Length:	2910
Score:	575.50	Matches:	237
Percent Similarity:	39.0%	Conservative:	159
Best Local Similarity:	23.3%	Mismatches:	337
Query Match:	10.8%	Indels:	283
DB:	10	Gaps:	41

US-10-645-746-3 (1-1020) x ADC46702 (1-2910)

Db 2017 GATGCTATTAGAAACATACCTCTTATTACTGATCGTCAACCATCATCATCGGTGCT 2076
 Qy 718 AspValThrHisProThrSerTySerGlyIleAspTySerIleAlaValAla 737
 Db 2077 GATGTGACTCACCCACAGCTGGAGAGGACTCAAGTCTCTTATTGTCTGTTGTGCC 2136
 Qy 738 SerIleAsn---ProGlyGlyThrIleTyArgSerMetIleValThrGln----- 753
 Db 2137 TCTATGAGTGCCTGAGATAAACAATACCGAGGATTTGTTCTGCTCAAGCTCATVAGG 2196
 Qy 754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
 Db 2197 GAGAAATATTTCAGGACCTGTATAAGCTGTTCCAGGATCCACAACTGGGTAGTCCAC 2256
 Qy 766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
 Db 2257 -----TCTGTTTGATAGGGAACATTTTCATAGCATTCAGGAGAGCT----- 2298
 Qy 786 GluAsnAsnAspAsnAlaProAlaHisIleValValTyArgAspGlyValSerAsp 805
 Db 2299 -----ACAGGCCAGATACCTCAAGAGATCATCTTCTATCGTGCAGGATGAAGCGNA 2349
 Qy 806 SerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGln 825
 Db 2350 GGCAGCTTTAGTCAGGTTCTGCTACATGATGATGCTATCCCAAGGCTTGTAACTCT 2409
 Qy 826 PheMetSerGluArgAspGlyGluAspProGluProLysTyThrPheIleValIleGln 845
 Db 2410 CTCGAA-----GAGAAATTATGTTCTCGTGTACTTTTCGTTGTTGTCAG 2454
 Qy 846 LysArgHisAsnThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLys 865
 Db 2455 AAACGTCCACACACACGCTTTGTTCCCTGAGCAACACCGG-----AATCGT 2499
 Qy 866 AspLeuThrProAlaGluThrAspValAlaValAlaLysGlnTrpGluGluAsp 885
 Db 2500 GATATGACT----- 2508
 Qy 886 MetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLys 905
 Db 2509 -----GATAAGAGTGGCAATATTCAACCA-----GGTACTGCTGGGACACT 2550
 Qy 906 LeuIleValSerLysTyLysPheAspPhePheLeuAlaSerHisGlyValLeuGly 925
 Db 2551 AAATCTGTCCACCTAATGAATTGCACTTCTATTGAACAGCCATGCTGTTATTCAGGGA 2610
 Qy 926 ThrSerArgProGlyHisTyThrValMetTyAspAspLysGlyMetSerGlnAspGlu 945
 Db 2611 ACAGCAGGCGCGCACATTACCATGTTCTTCGATGAGAACCGTTTCCCGCTGATCAG 2670
 Qy 946 ValTyLysMetThrTyGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSer 965
 Db 2671 TTGCAAAATGCTCAACAACCTCTGCTACACGATGCGAGGTGTACAAATCTGTGTCA 2730
 Qy 966 LeuProValProValHisTyAlaHisLeuSerCysGluLysAlaLysGluLeuTyArg 985
 Db 2731 ATTGTGCCACCGCTACTACGCTCACTTGGCTGCTATTCGTCGCTCC-----CGC 2778
 Qy 986 ThrTyLysGluHisTyIleGlyAspTyAlaGlnProArgThrArg 1001
 Db 2779 TACTACATGAGAGTGATGTCGTGATGGAGGTTCGAGGTCGAGGTCAGG 2826

RESULT 10

ADO02338

ID ADO02338 standard; cDNA; 2910 BP.

XX AC ADO02338;

XX DT 01-JUL-2004 (first entry)

XX DE Thalearess transcription factor cDNA #376.

XX KW Thalearess; transcription factor; ss; gene; plant; transgenic;

KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
 KW phosphate limitation; potassium limitation; nitrogen limitation;
 KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
 KW flowering; inflorescence architectural change;
 KW meristem cell differentiation; phyllotaxy; apical dominance;
 KW trichome development; seed development; premature senescence;
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.

OS Arabidopsis thaliana.

XX US2004045049-A1.

XX 04-MAR-2004.

XX 10-APR-2003; 2003US-00412699.

XX 13-SEP-1999; 99US-00394519.

XX 21-JAN-2000; 2000US-00489376.

XX 17-FEB-2000; 2000US-00506720.

XX 22-MAR-2000; 2000US-00532591.

XX 22-MAR-2000; 2000US-00533029.

XX 22-MAR-2000; 2000US-00533030.

XX 22-MAR-2000; 2000US-00533392.

XX 06-APR-2000; 2000US-00533648.

XX 16-NOV-2000; 2000US-00713994.

XX 27-MAR-2001; 2001US-00819142.

XX 17-APR-2001; 2001US-00837444.

XX 30-JAN-2002; 2002US-00958131.

XX 14-JUN-2002; 2002US-00171468.

XX 09-AUG-2002; 2002US-00225066.

XX 09-AUG-2002; 2002US-00225067.

XX 09-AUG-2002; 2002US-00225068.

XX 17-DEC-2002; 2002US-0434166P.

XX 25-FEB-2003; 2003US-00377480.

(ZHAN/) ZHANG J.

(FROM/) FROMM M E.

(HEAR/) HEARD J E.

(RIEC/) RIECHMANN J L.

(ADAM/) ADAM L J.

(BROU/) BROUN P E.

(PINE/) PINEDA O.

(REUB/) REUBER T L.

(KEDD/) KEDDIE J S.

(YUGG/) YU G.

(JIAN/) JIANG C.

(SAMA/) SAMAHA R S.

(PILG/) PILGRIM M L.

(CREE/) CREELMAN R A.

(DUBE/) DUBELL A N.

(RATC/) RATCLIFFE O.

(KUMI/) KUMIMOTO R.

(SHER/) SHERMAN B K.

Zhang J, Fromm ME, Heard JB, Riechmann JL, Adam LJ, Broun PE;

Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;

Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;

Sherman BK;

WPI; 2004-225755/21.

P-PSDB; ADO02339.

XX New transgenic plant, useful in developing phenotypes with altered or

XX improved characteristics or traits.

XX Claim 1; SEQ ID NO 751; 213pp; English.

XX The invention relates to a transgenic plant comprises a recombinant

XX polynucleotide having a polynucleotide sequence or its complementary

XX sequence comprising a sequence encoding a polypeptide, that initiates

XX transcription (i.e. a transcription factor) from Arabidopsis, Soybean,

XX

CC Rice, Rape or Corn, comprising any of the sequences appearing as AD001588
 CC -AD003527 or AD003530-AD003559. Also included are using a transgenic
 CC plant to grow a progeny plant, an expression cassette (comprising a
 CC constitutive, inducible or tissue-specific promoter and a recombinant
 CC polynucleotide described above), a host cell comprising the expression
 CC cassette, producing a modified plant having a modified trait, identifying
 CC a factor that is modulated by or interacts with a polypeptide encoded by
 CC the polynucleotide sequence and identifying at least one downstream
 CC polynucleotide sequence that is subject to a regulatory effect of any of
 CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed development, altered cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenly lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenly lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence encodes a
 CC thalecress transcription factor of the invention.

XX Sequence 2910 BP; 825 A; 590 C; 724 G; 771 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,24e-43 Length: 2910
 Score: 575.50 Matches: 237
 Percent Similarity: 39.0% Conservative: 159
 Best Local Similarity: 23.3% Mismatches: 337
 Query Match: 10.8% Indels: 283
 DB: 12 Gaps: 41

US-10-645-746-3 (1-1020) x AD002338 (1-2910)

QY 38 LysLysValLeuLeuValAenTrpPheLysPheSerSerLysIleTyrAspArgGlu 57
 DB 472 AGAAGATCATGGTTCGTGGGAT-----CATTTCTGTTCAGTTCGTATCGTAT 525
 QY 58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAenArgLysProGlyLysPro 77
 DB 526 CTCACCATACCATGTTTCGATCAATCTCGAGGTTATATCAAG----- 570
 QY 78 PheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
 DB 570 ----- 570
 QY 98 ArgHisGluLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp 117
 DB 1417 CAAAGATACACCAAAAGGCTCAATGAGAACGAAGTGCATGCTCAATGCTAAAGACTCCTCGC 1476

DB 570 ----- 570
 QY 118 ThrValTyrSerValCysArgLeuAenThrValThrSerLysMetLeuValSerGluLys 137
 DB 571 -----ACAGTGAACAGAAACGTGATCGAACTTCTGGTTAAGAAAT--- 609
 QY 138 ValValLysLysAspSerGlu-----LysLysAspGluLysAspLeuGluLysLys 154
 DB 610 -----TATAAAGATCTCTCCTGGGAGGAAGTACCAGCGGTATGAT---GGAAGGAAA 660
 QY 155 IleLeuTyrThrMet-----IleLeuThrTyrArgLysLysPheHisLeuAenPheSer 172
 DB 661 AGCCTCTATACCTGCTGGTCCATTACCTTTTGAGCTCGAAAGAGTTTGTGTGTAATCTGGCG 720
 QY 173 ArgGluAenProGluLysAspGluGluAaenArgSerTyrLysPhe---LeuLysAen 191
 DB 721 GAGAAAGAGCTGACGGTCTCTCTGGGAGGACAGACCGCTTTTAAAGTTCTGTGTAAAGAT 780
 QY 192 Val-----MetThrGlnLysValArgTyrAla 200
 DB 781 GTGACAAGCAGTGTCTTTATCAGTTGTCACAGCTTCCTTGATCGTGAAGCAAGAGAGCT 840
 QY 201 PropheValAsnGluGluLysValGlnPheAlaLysAenPheValTyrAspAenAsn 220
 DB 841 CCATAT-----GATACTATCCAGTGTGATGTTGTTCTTAGGGATAAGCCCTCTAAT 894
 QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeu--- 239
 DB 895 GATTATGCTCTCTGTTGGGAGGCTCTTT-----TTCACACTAGTTGTTGGGA 939
 QY 240 -----GluValAlaProArgIleGluAlaTrpPheGlyIle 251
 DB 940 AAGGACGCAAGAGATGTTAGGGGTGAGCTTGAGATGGTATTGAGTACTGGAGAGTTAT 999
 QY 252 TyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAenPheAlaIleValAsp 271
 DB 1000 TTCCAAAGTCTAAGG---CTGACTCAGATGGGTTTGTCTCTGAACACTTACGCTTTCAGCA 1056
 QY 272 LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuValAsp 291
 DB 1057 AGATCATTTTAT----- 1071
 QY 292 ProGlnSerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
 DB 1072 CGGATTGTTGTCAGTACTTATTAGCAAGTTCTTGAATATAAGGACTTA-----AAC 1125
 QY 312 MetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAenLeuLysLeu 331
 DB 1126 AGGCCACTTAGAGACTCAGATCGACTTAAGGTGAAGAAAGTTTTCAGGACACTGAAAGTT 1185
 QY 332 LysCysAlaGluValTrpAspAsnGluMetSerArg----- 343
 DB 1186 AAGTTGCTTAC---TGAACCGGCACAAAGAGTGCATAAATTAGTGGGATTTCTAGCTCTA 1242
 QY 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyr 362
 DB 1243 CCCATCAGGAGCTAAGGTTCACTCTGGAGGCAAAATCAGAGACGGTGTTCATAT 1302
 QY 363 LysValThrGlyLysSerAspArgGlyArgAenAlaLysLysTyrAspThrThrLeuPhe 382
 DB 1303 -----TTTGTGAAAAATAATAAT----- 1320
 QY 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
 DB 1321 -----TATAGA-----GTGAATACCAGGCTCTACCTGCTCTTCAACA 1359
 QY 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422
 DB 1360 GGGAGTGACACAAAGCCGCTTACCTACCAATGGAGCTCTGCCAAATTCACGAA---GGG 1416
 QY 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
 DB 1417 CAAAGATACACCAAAAGGCTCAATGAGAACGAAGTGCATGCTCAATGCTAAAGACTCCTCGC 1476

XX WPI: 2004-122948/12.
 DR P-PSDB; ADJ94698.
 XX New single-stranded RNA molecule having a length from 14-50 nucleotides,
 PT useful for preventing or treating tumor, inflammatory, infectious, e.g.
 PT viral infections, degenerative and autoimmune diseases.
 XX
 XX Example; Fig 16; 73pp; English.
 XX
 CC The invention relates to a novel single-stranded RNA molecule having a
 CC length from 14-50 nucleotides where at least 14-20 of the 5'-most
 CC nucleotides are substantially complementary to a target transcript. The
 CC RNA molecule of the invention demonstrates cytostatic, antiinflammatory,
 CC virucide and immunosuppressive activities and may be useful for
 CC inhibiting the expression of a target gene in vitro or in vivo,
 CC overexpression for preventing or treating diseases associated with the
 CC overexpression of at least one target transcript. The diseases may be
 CC selected from tumour diseases, inflammatory diseases, infectious diseases,
 CC such as viral infections, degenerative diseases and autoimmune diseases.
 CC Furthermore, the molecules of the invention may be utilised during gene
 CC therapy. The current sequence is that of the human eukaryotic translation
 CC initiation factor 2C2 (eIF2C2) cDNA of the invention.
 XX
 SQ Sequence 2580 BP; 625 A; 750 C; 706 G; 499 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,48e-42 Length: 2580
 Score: 566.00 Matches: 239
 Percent Similarity: 40.0% Conservative: 159
 Best Local Similarity: 24.0% Mismatches: 385
 Query Match: 10.6% Indels: 214
 DB: 12 Gaps: 38

US-10-645-746-3 (1-1020) x ADJ94704 (1-2580)

QY 19 ProGluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
 DB 52 CCATCCAGGATATGCTTCAAGCTCCACCTAGACCCGACTTTGGGACCTCCGGAGA 111
 QY 39 LysValLeuLeuValAenTrpPheLysPheSerSerLysLysLysLysLysLysLys 58
 DB 112 ACAATCAAAATTACAGGCCAATTTCTCGAATGGACATCCCAAAATTCACATCTATCAT 171
 QY 59 TyrGluTyrGluValLysMetThrLysGluValLeuAenArgLysProPhe 78
 DB 172 TATGATTCGATATCAAGCCAGACAGAG-----GTCCAG 201
 QY 79 ProLysLysThrGluLeuProLeuProLeuProLeuProLeuProLeuProLeuProLeu 98
 DB 202 CCGAGGAGATTAAACAGGGAATCGTGGAAACACATG-----GTCCAG 243
 QY 99 HisGluLysGlnThrAspPheLysLysLysLysLysLysLysLysLysLysLysLys 116
 DB 244 CACTTTAAACACAG-----ATCTTTGGGGATCGGAAGCCCGTGTTCACGCGCAGG 294
 QY 117 AspThrValTyrSerValCysArgLeuAenThrValThrSerLysMetLeuValSerGlu 136
 DB 295 AAGATCTATACACGCCATGCCCCCTTCGATTGGGAGGACAGGTG-----GAG 345
 QY 137 LysValValLysLysAspSerGluLysLysAspGluLysAspGluLysLysLysLysLys 156
 DB 346 CTGGAGGTCCAGTCGAGGAGAGGCAAGGATCGC-----ATCTTCAAGGTGTCCATC 399
 QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAenPheSerArgGluAenPro 176
 DB 400 AAGTGGGTCTCTCGGTAGCTTGCAGGCGTTACACGATGCATCTTCAGGGCGGTGCC 459
 QY 177 GluLysAspGluGluAlaAenArgSerTyrLysPheLeuLysAenValMetThrGln--- 195
 DB 460 ACGGTCCCTTTAG-----ACGATCCAGGCCCTGGAGCTGGTATGAGGACTTG 510
 QY 196 ---LysValArgTyrAlaProPheValAenGluGluLysValGlnPheAlaLysAen 214

511 CCATCCATGAGGTACACCCCGTG----- 534
 215 PheValTyrAspAenAenSerIleLeuArgValProGluSerPheHisAspProAenArg 234
 535 -----GGCCGCTCTTCTTCCACCGCTCGAAGGCTGCTCTAACCT----- 576
 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTirPheGlyLysIleGly 254
 577 -----CTTGGCGGGCCGAGAGTGTGTTTGGCTTCCATCAGTCC 618
 255 IleLysGluLeuPheAspGlyGluProValLeuAenPheAlaIleValAspLysLeuPhe 274
 619 GTCCGCGCTCTCTC-----TGGAAATGATGCTGAATATGATGTGTGCAGCAACAGCGTTT 675
 275 TyrAenAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuLeuLeuLeuLeuLeuSer 294
 676 TACAAGGCA-----CAGCCAGTAACTGAGTTTGTGTAAGTTTGGATTTTAAAGT 729
 295 CysAenAspValArgLysAspLysThrLysLeuMetAlaGlyLysMetThrIle 314
 730 ATTGAAGAACAA-----CAAAACCTCTG 753
 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAenLysLysLysLysLysLys 332
 754 ACAGATTCCCAAGGGTAAAGTTTACCAAGAAATTAAGGTTTAAAGGTGAGATAACG 813
 333 ---CysAlaGluValTrpAspAen-----GluMetSerArgLeuThrGluArg 347
 814 CACTGTGGCAGATGAAGAGGAGTACCGTGTCTCAATGTGACCGCGCGCGCCAGT 873
 348 HisLeuThrPheLeuAspLeuCysGluGluAenSerLeuValTyrLysValThrGlyLys 367
 874 CACCAAAATTC-----CCGCTGCAGCAGGAGCGGCGGACGCGTGGAGTGCAGC----- 924
 368 SerAspArgGlyArgAenAlaLysLysTyrAspThrThrLeuPheLysLysLysLysLys 387
 925 -----GTGCCCATGAT-----TTCAAG-----GACAGG 948
 388 AsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLys 407
 949 CACAAGTTGGTTCCTCGCTACCCCTCCATGTTTACAAGTCGACGAGCAGGACGAGAAA 1008
 408 GluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAen 427
 1009 CACACTACCTTCCCTCGAGGTCTGAACATTGTGGCAGGACAAAGATGATTTAAAAA 1068
 428 ArgIleAspLeuValMetGlnAspLysPheLysLysArgAlaThr---ArgLysProHis 446
 1069 TTAACGGAC-----AATCAGACCTCAACCATGATGATGATGATGATGATGATGATGAT 1122
 447 AspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAenPheSerSerGluLeu 466
 1123 GATCGCAGGAGAGAGATTAGCAATATGATGCGAAGTGCAGTTTCAACAGATCCA--- 1179
 467 AenPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLys 486
 1180 ---TAGTCCCGTGAATTTGGAATCATGTGTCAAAGATGATGATGATGATGATGATGATGAT 1236
 487 ValLeuLysGluProMetLeuValAenSerValAenGluGlnIleLysMetThrProVal 506
 1237 GTGTGTCAGCGCCCTCCATCTCTACGGGGGAGCAATAAAGCTATTTCGACCCCTGTC 1296
 507 -----IleArgGlyPheGlnGluLysGlnLeuAenVal---ValProGluLysGluLeu 523
 1297 CAGGCGCTGCGGACATCGGAAACAGCAGTTCCACCGGATCGAGATCAAGGTGTGG 1356
 524 CysCysAlaValPheValValAenGluThrAlaGlyAenProCysLeuGluAenAsp 543
 1357 GCATTGCGTGTCTCGCCCGCCAG-----CGCCAGTGCAGCGAAGTCCAT--- 1401
 544 ValValLysPheTyrThrGluLeuLeu-----GlyGly-CysLysPhe 557

Db 1402 ---CTGAAGTCCTTTCACAGACGCTCAGAAAGATCTCGAGAGCGCTGGCATGCCCATC 1458
 Qy 557 eArgGlyleAArgGlyleAlaAenGluAenAArgGlyleAlaGlnSerIle-----MetTy 575
 Db 1459 CAGGGCCAGCGCGTCTTCTGCAATACGCG-CAGGGGGCGGACAGCGTGGAGCCCATGTT 1517
 Qy 575 xAspAlaThrLysAenGluTyAlaPheTyrLysAenCysThrLysAenThrGlyIleG1 595
 Db 1518 CCGGCACCTGAGAACACGATGCG----- 1542
 Qy 595 yArgPheGluIleAlaAlaThrGluAlaLysAenMetPheGluArgLeuProAspLysG1 615
 Db 1542 ----- 1542
 Qy 615 uGlnLysValLeuMetPheIleIleIleSerLysAArgGlnLeuAenAlaTyArgLysPheVa 635
 Db 1543 -----GGCGTCAGCTGGTGGTGTCTATCTGCCCGCAAGACGCCCTGTACGCCGAGGT 1598
 Qy 635 lLysHisTyrCysAspHisThrIleGlyValAlaAenGlnHisIleThrSerGluThrVa 655
 Db 1599 CAAGCGGTGGAGACACGCTGCTGGGGATGGCCACGACGTGGTGCGAGATGAAGACGT 1658
 Qy 655 lThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysAArgIlePheTyrGlnI1 675
 Db 1659 G-----CAGAGACCAAGCCACAGACCTGTCCAACT 1691
 Qy 675 eAlaLeuLysIleAenAlaLysLeuGlyGlyIleAenGlnGluLeuAspTrpSerGluI1 695
 Db 1692 TTGCCTGAAGTCAACCTCAAGCTGGGAGCGGTGAACAACATCTG----- 1737
 Qy 695 eAlaGluIleSerProGluGluLysGluArgLysThrMetProLeuThrMetTyrVa 715
 Db 1738 -----CTGCCCCAGGGAGCGCGCGGTGTTCACAGCCCGTC-----ATCTTCT 1784
 Qy 715 lGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaVa 735
 Db 1785 GGGAGCAGACGTCACTACCCCGCGGGGATGGGAAGCCCTCCATTCGCCCGCT 1844
 Qy 735 lValAlaSerIleAenProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluG1 755
 Db 1845 GGTGGGCGAGCATGGAGCCCAATCGTACTGCGCCACCGTGGCGTGCGAGCAG-- 1902
 Qy 755 uCysAArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
 Db 1903 -----CACCGCGAGGAGATCATACAGACCTGCCCGC 1934
 Qy 775 aLysPheValLysLeuLeuArgGluPheAlaGluAenAsnAspAsnArgAlaProAlaHi 795
 Db 1935 CATGGTCCGCGACTCTCATCCAGTCTTACAGTCCACCGGCTTCAAG---CCCACCG 1991
 Qy 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG1 815
 Db 1992 CATCATCTTCTACCGGCGAGCGGTCTCTGAAGGCCAGTTCCAGCAGGTTCTCCACCACGA 2051
 Qy 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyLeuAspPr 835
 Db 2052 GTTGTGGCCATCCGTGAGCGCTGTATCAAGCTA-----GAAAGAGACTA 2096
 Qy 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAenThrArgLeuLeuArgAr 855
 Db 2097 CCAGCCCGGGATCATCTTATCTGTGTGCGAGAGAGGCACACCCCGGCTCTTC---TG 2153
 Qy 855 gMetGluLysAAspLysProValValAenLysAAspLeuThrProAlaGluThrAspValAl 875
 Db 2154 CACTGACAAAGACGAGCGGTGGGAAAGTGGAAACATTCACGCA----- 2199
 Qy 875 aValAlaAlaValLysGlnTrpGluLysMetLysGluSerLysGluThrGlyIleVa 895
 Db 2199 ----- 2199
 Qy 895 lAenProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
 Db 2200 -----GGCAGCATGTGGACAGCAAAATCACCCACCCCGGATTCGACTT 2246

Qy 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValme 935
 Db 2247 CTACTCTGTAGTCACTGCGCTGCAGGCGGACAGGAGGCTTCCGCACTATCACGCTCT 2306
 Qy 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955
 Db 2307 CTGGGACGACAACTGTTCTCTGATGAGCTGCAGATCCTAAACCTACACGCTGTGTCA 2366
 Qy 955 eLeuSerAlaArgCysArgLysProIleSerLeuProValHisTyrAlaHisLe 975
 Db 2367 CACTACGTGGCTGCACACGCTCCGTGTCTATCCAGCGCCAGCATACGTCACT 2426
 Qy 975 uSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
 Db 2427 GGTGGCCCTCCGGGCCAGG---TACCACCTGGTGATGAAGAACAT 2469
 RESULT 12
 ABQ99304
 ID ABQ99304 standard; cDNA; 3011 BP.
 XX
 AC ABQ99304;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 Human coding sequence SEQ ID 37.
 DB
 KW Human; expressed sequence tag; EST; chromosome 8q24;
 KW haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 WO200259260-A2.
 XX
 01-AUG-2002.
 XX
 16-NOV-2001; 2001WO-US042950.
 XX
 17-NOV-2000; 2000US-00714936.
 XX
 (HYSE-) HYSEQ INC.
 XX
 Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI; 2002-590824/63.
 DR N-PSDB; ABP64718.
 XX
 New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 XX
 Claim 1; SEQ ID NO 37; 394pp; English.
 PS
 CC The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation, the
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The

CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridization, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3011 BP; 732 A; 877 C; 805 G; 597 T; 0 U; 0 Other;

Alignment Scores: 1.84e-42 Length: 3011
 Pred. No.: 566.00 Matches: 239
 Score: 40.0% Conservatives: 159
 Percent Similarity: 24.0% Mismatches: 385
 Best Local Similarity: 10.6% Indels: 214
 Query Match: 6 Gaps: 38
 DB:

US-10-645-746-3 (1-1020) x ABQ99304 (1-3011)

Qy 19 ProGluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
 Db 198 CCATCTCAAGGATATGCTTCAAGCCTCCACCTAGACCGCAGCTTGGGACCTCCGGGAGA 257
 Qy 39 LysValLeuLeuValAlaAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyr 58
 Db 258 ACATCAAAATTACAGCCCAATTTCTTGAATGGACATCCCAAAATTGACATCTATCAT 317
 Qy 59 TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhe 78
 Db 318 TATGAATTCGATATCAAGCCAGCAGAGAG-----TGCC 347
 Qy 79 ProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArg 98
 Db 348 CCGAGGAGAGTTAAACAGGGAATCGTGGAAACATG-----GTCCAG 389
 Qy 99 HisGluLysLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLys 116
 Db 390 GACTTTAAACACAG-----ATCTTTGGGGATCGGAAGCCCGTGTTGACGGCAGG 440
 Qy 117 AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
 Db 441 AAGAAATCTATACAGCCATGCCCTTCGATTGGGAGGACACAGGTG-----GAG 491
 Qy 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysLysLeu 156
 Db 492 CTGGAGGTCAACGCTGCGAGCTTGCAGGCGTTACAGATGCATCTTCAGGGCGGTGCC 545
 Qy 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
 Db 546 AAGTGGGTCTCTGCGTGCAGCTTGCAGGCGTTACAGATGCATCTTCAGGGCGGTGCC 605
 Qy 177 GluLysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln--- 195
 Db 606 AGCGTCCCTTTTGAG-----ACGATCCAGGCCCTGGACGTGCTCATGAGGCATTTG 656
 Qy 196 ---LysValArgTyrAlaProPheValAlaGlnGluLysValGlnPheAlaLysAsn 214
 Db 657 CCATCCATGAGTACACCCCGTG----- 680
 Qy 215 PheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg 234
 Db 681 -----GGCGGTCTCTTCTTCCACCGGTCCGAAGGCTGCTCTAACCT----- 722
 Qy 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGly 254
 Db 723 -----CTTGGCGGGGCCGAGAAGTGTGTTGGCTTTCATCAGTCC 764

Qy 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
 Db 765 GTCCCGCTCTCTCTC---TGGAAATATGTGTAATTTGATGTGTCACAAACAGCGTT 821
 Qy 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSer 294
 Db 822 TACAAGGCA-----CAGCCAGTATTCAGATTTGTTGTGAAGTTTGTGATTTTAAAGT 875
 Qy 295 CysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
 Db 876 ATTGAAGAACA-----CAGCCAGTATTCAGATTTGTTGTGAAGTTTGTGATTTTAAAGT 899
 Qy 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLys----- 332
 Db 900 ACAGATTCCTCAAGAGGTAAAGTTTACCAAGAAATTAAGGTCTAAAGGTGGAGATAACG 959
 Qy 333 ---CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGluArg 347
 Db 960 CACTGTGGCAGATCAAGGNAAGTACCGTGTCTCAATGTGACCGCGCGGCCCGCACT 1019
 Qy 348 HisLeuThrPheLeuAspLysCysGluGluAsnSerLeuValTyrLysValThrGlyLys 367
 Db 1020 CACCAAAACATTC---CCGCTGCAGCAGGAGCGGCGCAGCGGTGGAGTGCAAG----- 1070
 Qy 368 SerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGlu 387
 Db 1071 -----GTGCCCGATAT-----TTCAG-----GACAG 1094
 Qy 388 AsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLys 407
 Db 1095 CACAGTTGGTTCGCGTACCCCACTCCCATGTTTACAAGTCGACAGGAGCAGAGAA 1154
 Qy 408 GluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsn 427
 Db 1155 CACACTACCTTCCCTCGAGGTCTGTAACTTGTGGCAGGACAAAGATGTATTAAAAA 1214
 Qy 428 ArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---ArgLysProHis 446
 Db 1215 TTAACGGAC-----AATCAGACCTCAACCATGATCAGAGCGCTGCTAGTGGCGGCC 1268
 Qy 447 AspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeu 466
 Db 1269 GATCGCAAGAGAGATTTAGCAAAATTGATCGCAAGTGCAGTTTCAACACAGATCCA--- 1325
 Qy 467 AsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLys 486
 Db 1326 ---TACGTCCGTGAATTTGGAATCATGTGTTAAAGATGAGATGACAGACGTGCTGGCGG 1382
 Qy 487 ValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal 506
 Db 1383 GTGTGTCAGCGCCCTCTCATCTCTACGGGGCAGAAATAAGCTATTTCGACCCCTGTCT 1442
 Qy 507 -----IleArgGlyPheGlnLysGlnLeuAsnVal---ValProGluLysGluLeu 523
 Db 1443 CAGGCGCTGCGACATCGGAACAAGCAGTTTCCACACGGGCATCGAGATCAAGGTGTGG 1502
 Qy 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
 Db 1503 GCCATTGGTGTCTCGCCCGCCAG-----CGCCAGTCACGGAAGTCCAT--- 1547
 Qy 544 ValValLysPheTyrThrGlnLeuIle-----GlyGly-CysLysPhe 557
 Db 1548 ---CTGAAGTCTTTCACAGCAGCTCAGAAAGATCTCGAGAGAGCGCGGATGCCCATC 1604
 Qy 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIle-----MetTy 575
 Db 1605 CAGGCCAGCGTGTCTTGTCAAAATACGCG-CAGGGGGCGGACAGCGTGGAGCCCATGTT 1663
 Qy 575 rAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlu 595
 Db 1664 CCGCACCTTGAAGAACACGTATGCG----- 1688

QY 595 YArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysG1 615
Db 1688 ----- 1688
QY 615 uGlnLysValLeuMetPheIleIleIleSerLysArgGlnLeuAenAlaTyrGlyPheVa 635
Db 1689 ---GGCTGCAGCTGGTGGTGGTCACTCTGCCCGCAAGACGCCCGGTAGCGCGAGGT 1744
QY 635 LysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVa 655
Db 1745 CAAGCGGTGGGAGACACGGTGTGGGATGCCACGCGAGTGGCTGCAGATGAAGAACGT 1804
QY 655 lThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnI1 675
Db 1805 G-----CAGAGACCCAGCCACAGACCCCTGTCCAACCT 1837
QY 675 eAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnLeuLeuAspTyrSerGluI1 695
Db 1838 CTGGCTGAAGATCAACGTCAAGCTGGGAGCGGTGAACAACATCTG----- 1883
QY 695 eAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrVa 715
Db 1884 -----CTGCCCCAGGCGAGCGCGCGGTGTTCAGCAGCGCCGTC--ATCTTTCT 1930
QY 715 lGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaIaVa 735
Db 1931 GGGAGCAGAGTCACTCAACCCCCCGCGGGATGGGAAGCCCTCCATTCGCGCGT 1990
QY 735 lValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluI1 755
Db 1991 GGTGGGCAGCATGAGCGCCACCCCAATCGTACTGGCCACCGTGGTGGCAGCAG-- 2048
QY 755 uCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
Db 2049 -----CACCGCGCAGGAGATCATACAAGACCTGGCGCG 2080
QY 775 aLysPheValLysLeuArgGluPheAlaGluAenAsnAspAsnArgAlaProAlaHi 795
Db 2081 CATGGTCCGCGAGCTCTCATCCAGTCTCAAGTCCACCGCTTCAAG---CCCACCGC 2137
QY 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG1 815
Db 2138 CATCATCTTCTACCGCGACGGTGTCTCTGAAGCCAGTTCACGAGCTTCTCACACGA 2197
QY 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
Db 2198 GTTGTCTGGCCATCGTGGAGCCCTGTATCAAGCTA-----GAAAAAGACTA 2242
QY 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgAr 855
Db 2243 CCAGCCCGGAGTCACTTCACTTCATCGTGGTGCAGAGAGGACCAACCCCGGCTCTTC--TG 2299
QY 855 gMetGluLysAspLysProValValAenLysAspLeuThrProAlaGluThrAspValAl 875
Db 2300 CACTGACAGAAGACGGGTGGGAAAGTGGAAACATTCAGCA----- 2345
QY 875 aValAlaAlaValLysGlnTrpGluLysAspMetLysGluSerLysGluThrGlyIleVa 895
Db 2345 ----- 2345
QY 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
Db 2346 -----GGCAGCAGCTGTGACACGAAATCAACCCACCCACCGAGTTCGACT 2392
QY 915 ePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
Db 2393 CTACCTGTGTAGTCAGCTGGCATCCAGGGACAGCAGCGCTTCGCACTATCACGTCCT 2452
QY 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955
Db 2453 CTGGGACGCAATCGTTTCTCTCTGATGAGTCGAGATCTCAACCTACCGCTGTGTCA 2512
QY 955 eLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHisLe 975

Db 2513 CACCTACGTGGCTGCACACGCTCCGTGTCCATCCAGCGCCAGCATACTACGCTCACCT 2572
QY 975 uSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db 2573 GGTGGCCTTCGGGCCAGG---TACCACCTGGTGGATGAAGNACAT 2615
RESULT 13
ABS78717
ID ABS78717 standard; cDNA; 3580 BP.
XX
AC ABS78717;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding NAAP7, from INCYTE no.1725129CB1.
XX
KW Human; ss; gene; nucleic acid associated protein; NAAP; cancer;
cell proliferative disease; cancer; atherosclerosis; hepatitis;
neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
epilepsy; developmental disorder; renal tubular acidosis; anaemia;
glaucoma; hypothyroidism; autoimmune disorder; AIDS;
inflammatory disorder; acquired immunodeficiency syndrome; allergy;
atopic dermatitis; arthritis; bacterial infection; viral infection;
parasitic infection; protozoal infection; fungal infection.
XX
OS Homo sapiens.
XX
PN WC0200272630-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003844.
XX
PR 09-FEB-2001; 2001US-0268118P.
PR 21-FEB-2001; 2001US-0270963P.
PR 22-FEB-2001; 2001US-0270858P.
PR 23-FEB-2001; 2001US-0271194P.
PR 07-MAR-2001; 2001US-0274071P.
PR 12-APR-2001; 2001US-0283496P.
PR 09-NOV-2001; 2001US-0344650P.
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
Gandhi AR, Ding L, Yue H, Gietzen KJ, Walia NK, Thangavelu K;
PI Elliott VS, Marquis JP;
XX
DR WPI; 2002-723320/78.
DR P-PSDB; ABG97473.
XX
PT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
treating and preventing diseases or conditions associated with the
aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
infections.
XX
PS Claim 5; Page 158-159; 162pp; English.
XX
CC The invention relates to an isolated polypeptide comprising one of 10
human nucleic acid associated protein (NAAP1-10), or a biologically
active or immunogenic fragment of the polypeptide, and their encoding
nucleic acid. Also included are a recombinant polynucleotide comprising a
promoter sequence operably linked to the polynucleotide, a cell
transformed with the recombinant polynucleotide, a transgenic organism
comprising the recombinant polynucleotide, an anti-NAAP antibody,
screening for a compound that is effective as an antagonist or modulator
of NAAP, generating an expression profile of a sample containing the
polynucleotides and an array comprising different nucleotide molecules
affixed on a solid substrate, nucleotide molecule comprises a first
oligonucleotide or polynucleotide sequence specifically hybridisable with
at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating

CC and preventing diseases or conditions associated with the decreased
 CC expression or overexpression of NAAp, such as cell proliferative diseases
 CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
 CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
 CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
 CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
 CC other diseases and disorders listed in the specification. These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of NAAp. The NAAp or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence encodes an NAAp protein
 XX
 SQ Sequence 3580 BP; 903 A; 953 C; 926 G; 797 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 2,34e-42 Length: 3580
 Score: 566.00 Matches: 239
 Percent Similarity: 40.0% Conservative: 159
 Best Local Similarity: 24.0% Mismatches: 385
 Query Match: 10.6% Indels: 214
 DB: 6 Gaps: 38

US-10-645-746-3 (1-1020) x ABS78717 (1-3580)

QY 19 ProGluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
 DB 95 CCATCCCAAGGATATGCTTCAAGCCTCCACCTAGACCGATCTGGGACCTCCGGGAGA 154
 QY 39 LysValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyr 58
 DB 155 ACAATCAAAATACAGGCCAATTTCTTGAATGGACATCCCAAAATTCGACATCATCAT 214
 QY 59 TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhe 78
 DB 215 TATGAATGGATATCAAGCCAGCAGAG-----TGC 244
 QY 79 ProLysLysThrGluIleProLysProAspArgAlaLysLeuPheTrpGlnHisLeuArg 98
 DB 245 CCAGGAGATTAACAGGGAATCGTGGACACATG-----GTCCAG 286
 QY 99 HisGluLysLysGlnThrAspPheLysLeuGluAspTyr-----ValPheAspGluLys 116
 DB 287 CACTTTAAACACAG-----ATCTTTGGGATCGGAAGCCCGTGTTCAGCGCAGG 337
 QY 117 AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
 DB 338 AAGAATCTATACACGCCATGCCCTTCGANTGGAGGCGGACAGGTG-----GAG 388
 QY 137 LysValValLysLysAspSerGluLysAspGluLysAspLeuGluLysLysIleLeu 156
 DB 389 CTGGAGGTACCGCTCCAGGAGAGCAGGATCGC-----ATCTTCAAGGTGTCATC 442
 QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
 DB 443 AAGTGGGTCTCGTGGAGTTCAGGCGTTACAGATGCATCTTCAGGCGCGGTGCC 502
 QY 177 GluLysAspGluLysAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln--- 195
 DB 503 AGCGTCCCTTTTTCAG-----ACGATCCAGCGCTCGAGCGTGTTCATGAGGCACCTTG 553
 QY 196 ---LysValArgTyrAlaProPheValAsnGluLysValGlnPheAlaLysAsn 214
 DB 554 CCATCCATGAGGTACACCCCGTG----- 577
 QY 215 PheValTyrAspAsnAsnSerLeuLeuArgValProGluSerPheHisAspProAsnArg 234
 DB 578 -----GGCGGCTCTTCTTCCCGGCTCGAAGGCTGCTCTAACCCCT----- 619

QY 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGly 254
 DB 620 -----CTTGGCGGGCGCGAGAGTGTGTTGGCTTCCATAGTC 661
 QY 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
 DB 662 GTCCGGCTTCTCTC---TGGNAATGATGCTGATATGATGTGTGCAGCAACAGCGTT 718
 QY 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuLeuLeuValAspProGlnSer 294
 DB 719 TACAAGGCA-----CAGCCAGTAATCGAGTTTGTGTAAGTTTGTGATTTTAAAGT 772
 QY 295 CysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
 DB 773 ATTGAAGACAA----- 796
 QY 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLys----- 332
 DB 797 ACAGATTCCCAAGGGTAAGTTTACCAAGAAATTAAGGTCTTAAGGTGGAGATAACG 856
 QY 333 ---CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGluArg 347
 DB 857 CACTGTGGGCAGATGAAGAGGAAGTACCGCGTCTGCAATGTGACCGCGCGCCGCACT 916
 QY 348 HisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLys 367
 DB 917 CACCAACATTC---CCGCTGCAGCAGGAGCGGCGGACGCGTGAGTGACG----- 967
 QY 368 SerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGlu 387
 DB 968 -----GTGCGCCAGTAT-----TTCAAG-----GACAGG 991
 QY 388 AsnLysLysPheIleGluPheProHisIleProLeuValLysValLysSerGlyAlaLys 407
 DB 992 CACAAGTTGGTCTCGCTACCCCACTCCCATGTTTACAAAGTCGGACAGGACGAGAA 1051
 QY 408 GluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsn 427
 DB 1052 CACACTTACCTTCCCTCGAGGTCTGTAACTTTGGCAGGCAAGATGATTATTAANA 1111
 QY 428 ArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---ArgLysProHis 446
 DB 1112 TTAACGGAC-----AATCAGACCTCAACCATGATCAGACGCTAGTGTGGCGGCC 1165
 QY 447 AspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeu 466
 DB 1166 GATCGCAAGAAGAGATTAGCAAAATTGATCGAAGTGCAAAGTTTCAACACAGATCCA--- 1222
 QY 467 AsnPheValGluArgPheGlyLysCysSerLysLeuMetIleGluCysProGlyLys 486
 DB 1223 ---TACGTCCGTGAATTTGGATCATGTGTCAAAAGATGATGATCAGACGCTGACTGGCGG 1279
 QY 487 ValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal 506
 DB 1280 GTGCTGCAGCGCCCTCCATCTCTACGGGGCAGGAATAAAGATTATTCGACCCCTGTC 1339
 QY 507 -----IleArgGlyPheGlnGluLysGlnLeuAsnVal---ValProGluLysGluLeu 523
 DB 1340 CAGGCGCTCTGGGACATCGGAACAGCAGATTCCACCGGCGATCGAGATCAAGGTGTGG 1399
 QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
 DB 1400 GCCATTGGTGTCTGCCCGCCAG-----CGCCAGTGCACGAAGTCCAT--- 1444
 QY 544 ValValLysPheTyrThrGluLeuIle-----GlyGly-CysLysPhe 557
 DB 1445 ---CTGAAGTCTTTCACAGCAGCTCAGAAAGATCTCGAGAGACGCTGGCATGCCATC 1501
 QY 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIle-----MetTy 575
 DB 1502 CAGGCGCCAGCGCTGCTTCTGCAAAATACCGC-CAGGGGGCGGACGCGTGGAGCCCATGTT 1560

QY 575 rAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCyseThrLeuAsnThrGlyIleG1 595
Db 1561 CCGCAGCTGAGAACAGTATGCG----- 1585
QY 595 yArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysG1 615
Db 1585 ----- 1585
QY 615 uGlnLysValLeuMetPheIleIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheVa 635
Db 1586 ---GGCCTGCAGCTGTGTGTGTCTATCGCCGCAAGACGCGCTAGCGCGAGGT 1641
QY 635 lLysHisTyrCyseAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVa 655
Db 1642 CAAGCGGTGGGAGACACGGTGTGGGGATGGCCAGCGAGTGTGCAGATGAAGAAGCT 1701
QY 655 lThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnI1 675
Db 1702 G-----CAGAGGACCACGCCACAGACCTGTCCAACT 1734
QY 675 eAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluI1 695
Db 1735 CTGCCTGAAGATCAACGCTCAAGCTGGGAGCGGTGAACAACATCTG----- 1780
QY 695 eAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrVa 715
Db 1781 -----CTGCCCGCAGGCGAGCGCGCGGTTCACAGCGCCGTC---ATCTTCT 1827
QY 715 lGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaVa 735
Db 1828 GGGAGCAGACGTCATCACCCCGCGCGGGATGGGAAGACCCCTCCATTGCGCGCT 1887
QY 735 lValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluI 755
Db 1888 GGTGGGCGAGCATGGAGCGCCACCCCAATCTACTGCGCCACCGTGGCGTGCAGCAG-- 1945
QY 755 uCyseArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
Db 1946 -----CACCGCAGGAGATCATCAAGACCTGCGCCG 1977
QY 775 aLysPheValLysLeuArgGluPheAlaGluAsnAspAsnArgAlaProAlaHi 795
Db 1978 CATGGTCCGCGAGCTCTCTATCCAGTCTTACAGTCCACCGCTTCAAG---CCACCCG 2034
QY 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG1 815
Db 2035 CATCATCTTCTACCGGCGAGCGTGTCTCTGAAGCCAGTTCACGAGGTTCTCCACCACGA 2094
QY 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
Db 2095 GTTGTGCGCCATCCGTGAGCGCTGTATCAAGCTA-----GAAAGAGCTA 2139
QY 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgAr 855
Db 2140 CCAGCCCGGATCACCTTCTATCTGTGTGCAGAGAGCCACACCGGCTTCTC---TG 2196
QY 855 gMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAl 875
Db 2197 CACTGACAGACGAGCGGTGGGAAAGTGGAAACATTCACGCA----- 2242
QY 875 aValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleVa 895
Db 2242 ----- 2242
QY 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
Db 2243 -----GGCAGCAGTGTGCACAGAAATCACCCACCCACCGAGTTCACCT 2289
QY 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
Db 2290 CTACCTGTGTAGTCAAGCTGGCATCCAGGCGGACAGCGGCTTCGACATCATCAGCTCT 2349
QY 935 tTyrAspAspLysGlyMetSerClnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955

Db 2350 CTGGGAGCAATCGTTTCTCTCTGATGAGCTGCAGATCCTAACCTACCACTGTGTGCA 2409
QY 955 eLeuSerAlaArgCyseArgLysProIleSerLeuProValProValHisTyrAlaHisLe 975
Db 2410 CACCTAGCTGGCTGCACAGCTCCGTGTCTATCCAGCGCCAGCATACTACGCTCACT 2469
QY 975 uSerCyseGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db 2470 GGTGGCCTCCGGGCCAGG---TACCACCTGGTGGATAAGGAACAT 2512
RESULT 14
AAH14510
ID AAH14510 standard; cDNA; 3996 BP.
XX
AC AAH14510;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12038.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isoqai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 12038; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent polynucleotides, all of which are used in the exemplification of the present invention


```

Db      1745  GTCCAACTCTGCTCAAGATCAATGTCAACTGGTGGCATTAACAGCATCTTA----- 1799
QY      692  pSerGluLeuAalaGluSerProGluGluLysGluArgGlyThrMetProLeuTh 712
Db      1800  -----GTCCCAACACAGCGCTCTGCCGTTTTCACACAGCGAGT-- 1838
QY      712  rMetTyrValGlyLeuAspValThrHisProThrSerTyrSerGlyLeuAspTyrSerIl 732
Db      1839  -ATATTCCTGGGAGCAGATGTTACACACCCCGAGCGGGATGGGAAAAAACCCTCTAT 1897
QY      732  eAlaAlaValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValTh 752
Db      1898  CACAGCAGTGTAGGCGATGATGATGCCACCCAGCCGATACGTCTACTGTGCGGT 1957
QY      752  rGlnGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIl 772
Db      1958  ACAG-----CGACCAGCGCAAGAGAT 1978
QY      772  eLeuGlu-----AlaLysPheValLys---LeuLeuArgGluPheAlaGluAsnAs 789
Db      1979  CATTGAAGACTTGTCTACATGTGTGCGTGAGCTCTCATCCNATTTTACAAGTCCACCG 2038
QY      789  pAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLe 809
Db      2039  TTTCAAG---CCTACCGCATCTCTCTCCGAGATGGGTGCTGAGAGCGAGTACC 2095
QY      809  uArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGl 829
Db      2096  CCAGATACTCCACTATGAGCTACTGGCCATTCTGTGATCGCTGCATCAAACTG----- 2147
QY      829  uArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAs 849
Db      2148  -----GAAAGGACTACACAGCTGGATCATCTATATTGTGTGGTGCAGAAACGCCATCA 2200
QY      849  nThrArgLeuLeuArgMetGluLysAspLysProValValAsnLysAspLeuThrPr 869
Db      2201  CACCCGCTTTTC---TGTCCTACAGAGATGGCGAATGGGAGAGTGGTAACATCCC 2257
QY      869  oAlaGluThrAspValAlaValAlaLysGlnTrpGluLysMetLysGluSe 889
Db      2258  AGCT----- 2261
QY      889  rLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSe 909
Db      2262  -----GGGACCACAGTGGGACACCAACATCACCCA 2290
QY      909  rLysTyrLysPheAspPheLeuAlaSerHisHisGlyValLeuGlyThrSerArgPr 929
Db      2291  CCATTTGATTTGATTTCTATCTGTGCACCCAGCGGATCCAGGGCCACAGCCGACC 2350
QY      929  oGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMe 949
Db      2351  ATCCCAATTAATATGTTCTTTGGGATGACAACCGTTTCACAGCAGATGAGCTCCAGATCCT 2410
QY      949  tThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValPr 969
Db      2411  GACGTACCAAGCTGTGCCACATTTACGTACGATGCACACGCTCTGTCTATCCACGACC 2470
QY      969  oValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGl 989
Db      2471  TGCTTACTATGCCCGCTGTGTGCTTCCGGGACCGA---TACCACCTGGTGGCAAGGA 2527
QY      989  uHisTyrIleGlyAsp 994
Db      2528  GCATGACAGTGGAGAG 2543

```

RESULT 15

ABSS1392

ID ABSS1392 standard; cDNA; 2955 BP.

XX AC

XX ABSS1392;

XX

```

DT      05-NOV-2002 (first entry)
XX      cDNA encoding larval viability associated protein #15.
XX      Fruit fly; larval viability; insecticidal activity; maize; wheat; oat;
XX      rye; sorghum; rice; barley; millet; turf; cotton; sugarcane; sugar beet;
XX      oilseed rape; soybean; vegetable crop; fruit; gene; ss.
XX      Drosophila melanogaster.
XX      WO200257455-A2.
XX      25-JUL-2002.
XX      18-JAN-2002; 2002WO-US001568.
XX      18-JAN-2001; 2001US-0262351P.
XX      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX      Stam L, Bachmann J, Broadus J, Kamdar KP;
XX      WPI; 2002-590746/63.
XX      P-PSDB; ABG70016.
XX      Identifying inhibitors of activity of proteins essential for Drosophila
XX      larval viability comprises expressing in a host a protein essential for
XX      larval activity and identifying compounds that inhibit or interact with
XX      the protein.
XX      Claim 1; Page 91-93; 169pp; English.
XX      The invention describes a method of identifying compounds that inhibit
XX      the activity of, or that interact with a protein essential for Drosophila
XX      larval viability comprising expressing in a recombinant host a DNA
XX      molecule to produce a protein essential for larval viability. The method
XX      is useful for identifying compounds with insecticidal activity. Compounds
XX      identified are useful as insecticides in crops such as maize, wheat,
XX      oats, rye, sorghum, rice, barley, millet, turf, cotton, sugarcane, sugar
XX      beet, oilseed rape, soybeans, vegetable crops and fruits. This sequence
XX      encodes a fruit fly larval viability associated protein
XX      Sequence 2955 BP; 651 A; 901 C; 807 G; 596 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2,48e-42      Length:      2955
Score:          564.50      Matches:      233
Percent Similarity: 40.4%      Conservative: 174
Best Local Similarity: 23.1%      Mismatches: 394
Query Match:    10.6%      Indels:      206
DB:              6      Gaps:      37

US-10-645-746-3 (1-1020) x ABSS1392 (1-2955)
QY      38  LysLysValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGlu 57
Db      436  CGCCCGATTGTGTGCGCCCAATCATCTCCAGGTGACA-----ATGCCCGGTGGC 486
QY      58  Tyr---TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLys 76
Db      487  TATGTGCATCATCATCATCATATACAGCCGACAGTGTCCGGAAGGTGACCGT 546
QY      77  ProPheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHis 96
Db      547  GAGATT-----ATCGAGACTATGGTGCATGCTATAGCAAGATCTTCGGA----- 591
QY      97  LeuArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLys 116
Db      592  -----GTGCTCAAGCCG---GTGTTTCGATGGTCCG 618
QY      117  AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
Db      619  AACATCTGTATACCGCGATCCCTGCCATTTGGCAACGAGCGTCTA-----GAG 669

```

```
QY 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuLysLysIleLeu 156
DB 670 CTGGAGGTTACTCTACCCGGCGAGGGCAAGATCGA----- 705
QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
DB 706 -----ATCTTCGGCTGACGATCAAGTGGCAGGCTCAGGTCTCGCTCTCAATTGG 756
QY 177 GluLysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLys 196
DB 757 GAGGAAGCTCTCGAAGCGCGC-----ACGGCGCAG 786
QY 197 ValArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheVal 216
DB 787 ATACCCATGATGCCATTTTGGCGTCGATGTGGTCATCGCGCATCTGCCAGCAGTACG 846
QY 217 TyrAsp-----AsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg 234
DB 847 TACAGCCAGTGGAGCTTCTTTCAGTTCCCGGAGGTTACTACCATCCC----- 900
QY 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTyrPheGlyIleTyrIleGly 254
DB 901 -----CTGGGTGTGACGCGAGGTTTGGTTCGGTTTCCATCAGAGC 942
QY 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
DB 943 GTAAAGCCCC---TCGCAGTGGAGAGTATGCTCAATATCGATGTCTGGGCCACCGCTTTC 999
QY 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuIleValAspProGlnSer 294
DB 1000 TACAAGGCT-----CAACAGTCATTGACTTCATGTGCGAGGTGCTGGACATTCGCGAC 1053
QY 295 CysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
DB 1054 ATCAACGAGCAG---CGCAACCGCTC----- 1077
QY 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLysCysAla 334
DB 1078 ACCGATTCGCAGCGCTCAAGTTACGAAGGAGATCAAGGGTTTGAAGATCGAGATCACC 1137
QY 335 GluValTyrAspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeu 354
DB 1138 CAC----- 1140
QY 355 CysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArgly----- 371
DB 1141 TCGCGCCAGATCGCTCGCAAGTATCGTGTGTGCAACAGTCACCTCGCGCGCCGCTCAGATG 1200
QY 372 -----ArgAsnAlaLysLysTyrAspThrThrLeuPheLysIle 384
DB 1201 CAATCATTTCCACTCGCTGAGAGAACGACAGACCGTAGAGTACCGCTGGCCCAAGTAC 1260
QY 385 Tyr---GluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysLys 403
DB 1261 TTCCTGGACAAGTACCGCATGAAATTCGCTACCGCATCTGCCCTCGCTGCGAGTTGGC 1320
QY 404 SerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGln 423
DB 1321 CAAGAGCACACACACTTACCTGCTCTAGAGGTGTGCAACATT---GTGGCCGACAG 1377
QY 424 ArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArg 443
DB 1378 CGGTGCATTTAAAGCTGACCATATGACAGCGTGCACATCATCAAGGCCACACAGCTCGT 1437
QY 444 LysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSer 463
DB 1438 TCAGCTCCGGATTCGTGAGCGTGAGATTAAACAATTGGTAAAGCGCGCGCATTTCAACAC 1497
QY 464 GluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCys 483
DB 1498 GAT-----TCGTATGTGCAAGAGTTTGGCTTGACCATCTCCAAATTCGATGATGAGGTA 1551
```

```
484 ProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMet 503
DB 1552 CGAGARCGGCTTCCCTCCTCCCAAGCTTCAGTATGGGGACGT----- 1596
QY 504 ThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValProGluLysGluLeu 523
DB 1597 -----GTGCTCTACCGGCTCACCGGCAGCAGCTG---TTCGCGCACAGAACAAGGTG 1647
QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
DB 1648 AGCTTGGCC-----TCGCCCAACACGAGGTGTATGGATATGCCA 1686
QY 544 ValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheArgGlyIleArgIleGly 563
DB 1687 GGCAAGCAGTCTTCTCCTCGGCTC-----GAGATCCGCATCTGG 1725
QY 564 AlaAsnGluAsnArgGlyAlaGln---SerIleMetTyrAspAlaThrLysAsnGluTyr 582
DB 1726 GCATCGCGCTGTTTTCGCCCACAGCGCAGCGTGCAGGATGCGCTGCGTAATTTCAAC 1785
QY 583 AlaPheTyrLysAsnCysThrLeuAsnThrGly-----IleGly----- 595
DB 1786 CAGCAGCTGCACAAGATCTCAACAGATCAGCGCATGCCGATAAATTTGGACAGCGCTGCTTC 1845
QY 596 ---ArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLys 614
DB 1846 TGTAGTAGCCACCGCGCGCGATCAAGTGGNACCATTGTTCCGTTACTTGAAGATCACC 1905
QY 615 GluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPhe 634
DB 1906 TTCCCGCGCTGCAACTCGTGTGTTGTGTCGCCGCAAGACTCCAGTATACGCCGAG 1965
QY 635 ValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThr 654
DB 1966 GTGAGCGCTAGGTGACACCGCTTCTGGGTATGGCCACCCAGCTGTGTGAGGCCAAGAAC 2025
QY 655 ValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGln 674
DB 2026 GTGAACAAGACA-----TCGCCACAGACGCTCTCTAAT 2058
QY 675 IleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpSerGlu 694
DB 2059 CTGTGCTGTGAAGATCAACGTCAGTTGGCGGCATCAATTTCAATTTCTG----- 2106
QY 695 IleAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyr 714
DB 2107 GTCCCTCCATTCGCCCAAGGTCTTCAATGAGCCG-----GTCACTCTT 2151
QY 715 ValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAla 734
DB 2152 TTGGTGGCGATGTGACACACCCACGCTGGCGACACACAAAGAACCATTCGATTCGCGCC 2211
QY 735 ValValAlaSerIleAsnProGlyThrIleTyrArgAsnMetIleValThrGlnGlu 754
DB 2212 GTGTGGGCTCCATGGATGCCCATCCATCGCGGTATGCCGCCACCGTTCGCGTACAGCAG 2271
QY 755 GluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu 774
DB 2272 -----CACCACAGAGAGATCATCCAGGAGCTGAGC 2301
QY 775 AlaLysPheValLysLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAla 794
DB 2302 AGCATGTGGCGAGCTGTTGATCATGTCTACAAGTCAGCGGGCGGTACAGCCCCAC 2361
QY 795 HisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAsp 814
DB 2362 CGCATCATACTCTATCGTGACGAGCTCCGAGGGACAAATTTCCACATCTCTCTGCAACAC 2421
QY 815 GluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAsp 834
DB 2422 GAATTGACCGCCATTCGCGAGCCCTGCATTAAAGCTA-----GAG 2460
QY 835 ProGlu-----ProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeu 852
```

[illegible]

Search completed: July 5, 2006, 22:39:32
Job time : 1412 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 6, 2006, 01:16:25 ; Search time 2698 Seconds
(without alignments)
6968.155 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 5349

Sequence: 1 MSSNFPLEKGFYHSLDPE.....RHEMFLQTNVKGPMSPA 1020

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10645746/runat_05072006_110158_5297/app_query.fasta_1
-DB=Published Applications.NA.Main -OPMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USER=US10645746 @CGN 1.1 1780 @runat_05072006_110158_5297 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOC=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.Main.*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	5349	100.0	3227	9	US-10-645-746-2

2	5349	100.0	3227	10	US-10-645-735-2	Sequence 2, Appli
3	5349	100.0	3227	16	US-11-144-985-2	Sequence 2, Appli
4	5001	93.5	3719	9	US-10-645-746-1	Sequence 1, Appli
5	5001	93.5	3719	10	US-10-645-735-1	Sequence 1, Appli
6	5001	93.5	3719	16	US-11-144-985-1	Sequence 1, Appli
7	577	10.8	7478	6	US-10-007-078-3	Sequence 3, Appli
8	577	10.8	7478	13	US-11-035-669-3	Sequence 3, Appli
9	575.5	10.8	2910	3	US-09-533-029-105	Sequence 105, App
10	575.5	10.8	2910	6	US-10-286-264-101	Sequence 101, App
11	575.5	10.8	2910	8	US-10-374-780A-2083	Sequence 2083, Ap
12	575.5	10.8	2910	8	US-10-412-699B-751	Sequence 751, App
13	566	10.6	3011	8	US-10-115-635-50	Sequence 50, Appl
14	566	10.6	3580	8	US-10-467-397-17	Sequence 17, Appl
15	564.5	10.6	3586	13	US-11-097-143-27926	Sequence 27926, A
16	564.5	10.6	3586	13	US-11-097-143-27929	Sequence 27929, A
17	564.5	10.6	3743	13	US-11-097-143-14510	Sequence 14510, A
18	556.5	10.4	2914	3	US-09-774-434-2	Sequence 2, Appli
19	556.5	10.4	3050	7	US-10-175-492-4	Sequence 4, Appli
20	556.5	10.4	3578	10	US-10-756-149-37	Sequence 37, Appl
21	550.5	10.3	3307	9	US-10-852-630A-4	Sequence 4, Appli
22	550.5	10.3	3325	10	US-10-483-505-66	Sequence 66, Appl
23	536.5	10.0	3509	8	US-10-437-963-89857	Sequence 89857, A
24	533.5	10.0	3509	8	US-10-437-963-18974	Sequence 18974, A
25	530.5	9.9	2815	7	US-10-269-909-20	Sequence 20, Appl
26	519.5	9.7	8465	13	US-11-097-143-27928	Sequence 27928, A
27	519.5	9.7	10149	13	US-11-097-143-27925	Sequence 27925, A
28	519.5	9.7	12558	13	US-11-097-143-14509	Sequence 14509, A
29	519	9.7	3147	8	US-10-374-780A-2085	Sequence 2085, Ap
30	509.5	9.5	2973	8	US-10-437-963-6231	Sequence 6231, Ap
31	507	9.5	3330	8	US-10-437-963-21837	Sequence 21837, A
32	507	9.5	3372	8	US-10-437-963-6229	Sequence 6229, Ap
33	502	9.4	3474	8	US-10-437-963-98130	Sequence 98130, A
34	501	9.4	2967	3	US-09-938-842A-432	Sequence 432, App
35	501	9.4	2967	3	US-09-938-842A-432	Sequence 432, App
36	499.5	9.3	2832	8	US-10-437-963-35635	Sequence 35635, A
37	499.5	9.3	3761	8	US-10-437-963-55972	Sequence 55972, A
38	496.5	9.3	3718	8	US-10-425-114-20232	Sequence 20232, A
39	495	9.3	2042	6	US-10-174-363-1	Sequence 1, Appli
40	495	9.3	2042	13	US-11-093-888-1	Sequence 1, Appli
41	494.5	9.2	3994	6	US-10-437-963-100929	Sequence 100929,
42	489	9.1	3705	13	US-10-174-363-53	Sequence 53, Appl
43	489	9.1	3705	13	US-11-093-888-53	Sequence 53, Appl
44	488	9.1	3323	7	US-10-225-066A-151	Sequence 151, App
45	488	9.1	3323	8	US-10-374-780A-237	Sequence 237, App

ALIGNMENTS

RESULT 1
US-10-645-746-2
; Sequence 2, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3227
; TYPE: DNA

; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21) ... (3080)
US-10-645-746-2

Alignment Scores:

Pred. No.: 0 Length: 3227
Score: 5349.00 Matches: 1020
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-645-746-3 (1-1020) x US-10-645-746-2 (1-3227)

QY 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
DB 21 ATGCTCTCGAATTTTCCCGAATTGGAAAAAGGATTTTATCGTCATTTCTCGATCCGGAG 80
QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysVal 40
DB 81 ATGAAATGGCTTTCGAGGCCCACTGGTAAATGCGCGCAAAATTTCTATGAGAAGAAGTA 140
QY 41 LeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGlu 60
DB 141 CTTCTTTTGGTAAATTTGGTTTCAAGTTCTCAGCAAAATTTTATCGATCGGGAATACTACGAG 200
QY 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
DB 201 TATGAAGTCAAAATGACAAGGAAGTATTGAATAGAAAAACCGAGMAAACCTTTCCCNAAA 260
QY 81 LysThrGluLysProLysProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGlu 100
DB 261 AAGACAGAAATTTCCAAATTTCCGATCGTGCANAACTCTTCTGCAACATCTTCGGCATGAG 320
QY 101 LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
DB 321 AAGACGACACAGATTTTATTTCTCGAAGACTATGTTTTTGTATGAAAAGGACACTGTTTAT 380
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValValLys 140
DB 381 AGTGTTTGTGCGACTGAACACTGTGCATCAAAATGCTGTTTCGGAGAAAGTAGTAAAA 440
QY 141 LysAspSerGluLysLysAspGluLysAspLeuGluLysIleLeuTyrThrMetIle 160
DB 441 AAGGATTCGGAGAAAAAAGATGAAAAGGATTTGGAGAAAAAATCTTATACAAATGATA 500
QY 161 LeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAspGlu 180
DB 501 CTTACCTATCGTAAAAAATTTCCCTGAACTTTTAGTCGAGAAAAATCCGGAAAAAAGCGAA 560
QY 181 GluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysValArgTyrAla 200
DB 561 GAAGCGAATCGGAGTTTACAAATTCCTGAAGAATGTTATGACCCAGAAAGTTCCGTACGCG 620
QY 201 ProPheValAsnGluLysValGlnPheAlaLysAsnPheValTyrAspAsn 220
DB 621 CCTTTTGTGAACGAGAGATTAAGTACAAATTCGCGAAAAATTTTGTGTACCATATAAT 680
QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeuGlu 240
DB 681 TCNAATTCGCGAGTTCTCGATCGTTTTCAGATCCAAACAGATTCGAACATCATTAGAA 740
QY 241 ValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyLysGluLeuPheAsp 260
DB 741 GTAGCACCAAGAATCGAAGCATGTTTGGAAATTTACATTTGAATCAAAAGATTTGTTCCGAT 800
QY 261 GlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMet 280
DB 801 GTGGAACCTGTGCTCAATTTTGGCAATTTGTGCAATAACTATTCTACAACTCACCAGAAATG 860
QY 281 SerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAspValArg 300

DB 861 TCTCTTCTGGATTATCTTCTCTAATTTGTCGACCCCGAGTCGTGTAAACGATGATGTACGA 920
QY 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro 320
DB 921 AAAGATCTTAAAACAAACTGATGCGGGGAAAAATGACAAATCAGACAAGCCGCGGCCCA 980
QY 321 ArgIleArgGlnLeuLeuGluAsnLeuLysCysAlaGluValTrpAspAsnGlu 340
DB 981 AGAATTCGACCAATTTATGGAAATTTGAAGCTGGAATGCGCAAGATTTGGGATATACGAA 1040
QY 341 MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuLysGluGluAsnSerLeu 360
DB 1041 ATGTCGAGATTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGGAAAACTCTCTT 1100
QY 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThr 380
DB 1101 GTTTATAAGTCACTGGTAAATCGCACAGAGAGAGAAATGCAAAAAAGTACGATATACTA 1160
QY 381 LeuPheLysIleTyrGluGluAsnLysPheIleGluPheProHisLeuProLeuVal 400
DB 1161 TTGTTCAAAATCTATGAGGAAAAAANAAGTTTCATTTGAGTTTCCCCACCTACACTAGTC 1220
QY 401 LysValLysSerGlyValLysGluTyrAlaValProMetGluHisLeuGluValHisGlu 420
DB 1221 AAAGTTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATGGAACATCTTGAAGTTCATGAG 1280
QY 421 LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLysArg 440
DB 1281 AAGCACAAAGATACAGAATCGAATTTGATCTGGTGATGCAAGCAAGTTTCTAAAGCGA 1340
QY 441 AlaThrArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAsp 460
DB 1341 GCTACACGAAAACCTCACGACTACAAAGAAAATACCTTAAAAATGCTGAAAGAAATGGAT 1400
QY 461 PheSerSerGluGluLeuAsnPheValGluArgPheGlyLeuLysSerLysLeuGlnMet 480
DB 1401 TTTCTCTCTGAAGAGCTAAATTTTCTGAAAGATTTGGATTTATGCTCCAAAATTCAGATG 1460
QY 481 IleGluCysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGln 500
DB 1461 ATCGAATGTCCAGGAAGGTTTGAAGAGCCCAATGCTTGTGAATAGTGTAAATGAACAA 1520
QY 501 IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGlu 520
DB 1521 ATTAATAATGACACCACTGATTTCTGCGATTTTCAAGAAAACAATTTGAATGCTGTTCCGAA 1580
QY 521 LysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGlu 540
DB 1581 AAAGAACTTTGCTGTCTGTTTTTGTAGTCAACGAAACAGCGGGAATCCATGCTTAGAA 1640
QY 541 GluAsnAspValValLysPheTyrThrGluLeuIleGlyCysLysPheArgGlyIle 560
DB 1641 GAGAACGACGTTGTTAAGTTCTACACCGAACTTAATTTGGTGGTTGCAAGTTCCGTGAATA 1700
QY 561 ArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsn 580
DB 1701 CGAATTTGGTGCATGAAAAACAGAGAGCGCAATCTATTATGTACGACCGCGAGAAAAAT 1760
QY 581 GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAla 600
DB 1761 GAATATGCTCTTCAAAAAATTTGTACACTAAATACCGGAATCCGTAGATTTGAAATAGCC 1820
QY 601 AlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMet 620
DB 1821 GCACAGAGCGGGAAGAAATATGTTTGAACGCTTCTCCCGATAAAGAACAAAGTCTTAATG 1880
QY 621 PheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAsp 640
DB 1881 TTCATTATCAATTTCCAAACGACAACTGAATGCTTACGTTTTTGTGAACATATTATTCGAT 1940
QY 641 HisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAla 660

```
Db 1941 CACACCATCGGTGTAGCTAATCAGCATATTACTTCTGAAACAGTCACAAAAGCTTTGGCA 2000
Qy 661 SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsn 680
Db 2001 TCACTAAGGCACGAGAAAGGATCAAAACGAATTTTCTATCAAAATTCGATTGAAATCAAC 2060
Qy 681 AlalysLeuGlyGlyIleAenGlnGluLeuAspTrpSerGluIleAlaGluIleSerPro 700
Db 2061 GCGAAATTAGAGGTATTAAACGAGGCTTGACTGGTCAGAAATTCGAGAAATATCACCA 2120
Qy 701 GluGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyIleAspValThr 720
Db 2121 GAAGAAAAGAAAGACCGAAACCAATGCCATTATGATGTATGTGGAAATGATGAAT 2180
Qy 721 HisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAlaSerIleAsn 740
Db 2181 CATCCAACTCTACACAGTGAATTTGATTTATATAGCGGTGTAGTAGCGAGTATCAAT 2240
Qy 741 ProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysArgProGlyGlu 760
Db 2241 CCAGGTGGAACTATCTATCGAAATATGATTTGTGACTCAAGAAGAAATGCTGTCGGGTGAG 2300
Qy 761 ArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeu 780
Db 2301 CGTGCAGTGGCTCATGGACGGGAAGAAACAGATATTTTGGAGCAAGTTCTGGAAATG 2360
Qy 781 LeuArgGluPheAlaGluAenAsnAspAsnArgAlaProAlaHisIleValValTyrArg 800
Db 2361 CTCAGAGAAATTCGAGAAAACACGACAAATCGAGCACCGCGCATATTTAGTCTATCGA 2420
Qy 801 AspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLys 820
Db 2421 GACGGAGTAGCGAATTCGGAGATGCTACGTGTGTAGTCATGATGAGCTTCGATCTTTAAA 2480
Qy 821 SerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThr 840
Db 2481 AGCGAGTAAACAAATTCATGTCGAAACGGATGGAGAAATCCAGAGCCGAAAGTACAG 2540
Qy 841 PheIleValIleGlnLysArgHisAenThrArgLeuLeuArgMetGluLysAspLys 860
Db 2541 TTCATTGTGATTTCAGAAAAGACACAAATACACGATTGCTTCGAAGAAATGGAAAAGATAAG 2600
Qy 861 ProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys 880
Db 2601 CCAGTGGTCAATAAAGATCTTACTCTCTGCTGAAACAGATGCTGCTGTGTCTGTAAA 2660
Qy 881 GlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleValAsnProSerSerGly 900
Db 2661 CANTGGAGGAGGATATGAAGAAGAAAGCAAGAACTGGAAATTTGTGAACCCATCATCCGA 2720
Qy 901 ThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHis 920
Db 2721 ACAACTGTGATAAACTTATCGTTTCGAATAACAATTCGATTTTCTTTGGCATCTCAT 2780
Qy 921 HisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGly 940
Db 2781 CATGGTGTCTCTGGTATCATCTGCTCCAGGACATATACATGTTATGTATGACGATAAGGA 2840
Qy 941 MetSerGluAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCys 960
Db 2841 ATGAGCCAGATGAGTCTATAAAGTAAAGACCTACGAGCTTCGTTCTCTCTGCTAGATG 2900
Qy 961 ArgLysProIleSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAla 980
Db 2901 CGAAAACCCATCTCGTTGCTGTTCCGGTTTCATTATGCTCATTTATCATGTGAAAAGCG 2960
Qy 981 LysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThr 1000
Db 2961 AAAGAGCTTTATCGAACTTTACAAGGAACATTAATCATCGGTGACTATGCACAGCCCGACT 3020
Qy 1001 ArgHisGluMetGluHisPheLeuGlnThrAsnValLysTyrProGlyMetSerPheAla 1020
Db 3021 CGACACGAATGGAACATTTTCTCCAAACTTAACGTGAAGTACCCTGGAAATGCTGTCGCA 3080
```

RESULT 2

```
US-10-645-735-2
; Sequence 2, Application US/10645735
; Publication No. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UMG-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(3080)
US-10-645-735-2
Alignment Scores:
Pred. No.: 0 Length: 3227
Score: 5349.00 Matches: 1020
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-645-746-3 (1-1020) x US-10-645-735-2 (1-3227)
```

```
Qy 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
Db 21 ATGTCCTCGAAATTTCCCGAAATTTGGAAAAAGGATTTTATCGTCATTCTCTCGATCCGGAG 80
Qy 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysVal 40
Db 81 ATGAATGGCTTCGGAGGCCCTGGTAAATGCGAGCGCAAAATTTATGAGAAAGAAAGTA 140
Qy 41 LeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGlu 60
Db 141 CTTCTTTTGGTAAATTTGGTTCAAGTTCTCCAGCAAAATTTACGATCGGGAATACTACGAG 200
Qy 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
Db 201 TATGAAGTGAATGCAAGGAAGTATTGAATAGAAAAACAGGAAAAACCTTTCCCAAAA 260
Qy 81 LysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGlu 100
Db 261 AAGACAGAAATTCCAATTCGGATCGTGGCAAAACTCTTCTGGCAACATCTTCGGCATGAG 320
Qy 101 LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
Db 321 AAGAACGACAGACAGATTTTATTTCTCGAAGACTATGTTTTTTCATGAAAGGACACTGTTTAT 380
Qy 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValValLys 140
Db 381 AGTGTGTCGACTGAACACTGTGCACATCAAAATGCTGCTTTTCGGAGAAAGTAGTAAAA 440
Qy 141 LysAspSerGluLysLysAspGluLysAspGluLysLysIleLeuTyrThrMetIle 160
Db 441 AAGGATTCGAGAAAAGAAAGATGAAAGGATTTGAGAGAAAATAATCTTTATACAAATGATA 500
```

QY	161	LeuThrTyrArgLysLysPheHisLeuAenPheSerArgGluAenProGluLysAspGlu	180	QY	521	LysGluLeuCysCysAlaValPheValValAenGluThrAlaGlyAenProCysLeuGlu	540
DB	501	CTTACCTATCGTAAATAAATTTTACCTCGAACTTTAGTCGAGAAATCCGAAAAAGACGAA	560	DB	1581	AAAGAACITTTGCTGCTGCTGTTTTTTGTAGTCAACGAAACACGCGGAAATCCATGCTTAGAA	1640
QY	181	GluAlaAenArgSerTyrLysPheLeuLysAenValMetThrGlnLysValArgTyrAla	200	QY	541	GluAenAspValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheArgGlyIle	560
DB	561	GAAGCGAATCGGAGTTACAAATTCCTGAAGAATGTTATGACCCAGAAAGTTGCTACGGG	620	DB	1641	GAGAACGACGTTGTTAAGTTCTACACCGAACTAAATTTGGTGGTTGCAAGTTCCGTGGAATA	1700
QY	201	ProPheValAenGluGluLysValGlnPheAlaLysAenPheValTyrAspAenAen	220	QY	561	ArgIleGlyAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAen	580
DB	621	CCTTTTGTGAACGAGGAGATTAAAGTACAAATTCGCGAAAAATTTTGTGTACGATAATAAT	680	DB	1701	CGAATTTGGTCCAAATGAAGAACAGAGAGCGCAATCTATTATGACGACCGACGAAAAAT	1760
QY	221	SerIleLeuArgValProGluSerPheHisAspProAenArgPheGluGlnSerLeuGlu	240	QY	581	GluTyrAlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyVargPheGluIleAla	600
DB	681	TCAAITTCGCGAGTTCCTGAATCGTTTACAGATCCAAACAGATTCGAATCAACATATTAGAA	740	DB	1761	GAATATGCTTCTACAAAAATTTGTACACTAAATACCGAATCGGTAGATTGAAATAGCC	1820
QY	241	ValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGluLeuPheAsp	260	QY	601	AlaThrGluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMet	620
DB	741	GTAGCACCAAGATCGAAGCATGTTTGGAAATTTACATTTGGAATCAAGAAATTTGTCGAT	800	DB	1821	GCAACAGAAAGCGAAGAAATATGTTTGAACGCTCTCCCGATAAAGAACAAAAAGTCTTAATG	1880
QY	261	GlyGluProValLeuAenPheAlaIleValAspLysLeuPheTyrAenAlaProLysMet	280	QY	621	PheIleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAsp	640
DB	801	GGTGAACCTGTCTCAATTTTGCATTTGTCGATAAATCTATTCTACAAATGCACCGAAAAATG	860	DB	1881	TTCAATATCATTTCCAAACGACAACTGAATGCTTACGCTTTTGTGAAACATTTATTCGAT	1940
QY	281	SerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAenAspValArg	300	QY	641	HisThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAla	660
DB	861	TCTCTTCTGGATATCTTCTCTAAATTTGTCGACCCCGCTGCTGTAACGATGATGTACGA	920	DB	1941	CACACATCGGTAGCTTAATCAGCATATTACTTCTGAAACAGTCAACAAAGCTTTGGCA	2000
QY	301	LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro	320	QY	661	SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAen	680
DB	921	AAAGATCTTAAACAAAACTGATGGCGGGAATAATGACAAATCAGACACCGCGCGGCA	980	DB	2001	TCACCTAGGACGACGAGAAAGGATCAAAACGAAATTTTCTATCAAAATGTCATTGAAAAATCAAC	2060
QY	321	ArgIleArgGlnLeuLeuGluAenLeuLysLeuLysCysAlaGlnValTrpAspAenGlu	340	QY	681	AlaLysLeuGlyGlyIleAenGlnLeuAenTrpSerGluIleAlaGluIleSerPro	700
DB	981	AGAAATTCGACAAATATTGGAATAATTTGAAGCTGAAATGCGCAGAAAGTTTGGGATAACGAA	1040	DB	2061	CGCAATTTAGGAGGTATTAAACAGAGCTTGACTGGTCAGAAATTCGAGAAATATCACCA	2120
QY	341	MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAenSerLeu	360	QY	701	GluGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyIleAspValThr	720
DB	1041	ATGTCGAGATTGACAGACGACATCTGACATTTCTAGATTTTGTGCGAGGAAATCTCTTT	1100	DB	2121	GAAGAAAAAGAAAGACGGAACAAATGCCATTAATATGATGTTGGAATGATGTAAT	2180
QY	361	ValTyrLysValThrGlyLysSerAspArgGlyArgAenAlaLysLysTyrAspThrThr	380	QY	721	HisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlaSerIleAen	740
DB	1101	GTTTATAAGTCACCTGGTAAATCGACACAGAGAAATGCAAAAAAGTACGATACATCA	1160	DB	2181	CATCAACCTCTACAGTGGAAATTCATTTCTATAGCGGCTGTAGTAGCGAGTATCAAT	2240
QY	381	LeuPheLysIleTyrGluGluAenLysLysPheIleGluPheProHisLeuProLeuVal	400	QY	741	ProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluGluCysArgProGlyGlu	760
DB	1161	TTGTTCAAAATCTATGAGGAAACAAAAAGTTTCATTGAGTTTCCCCCACCCTACCACTAGTC	1220	DB	2241	CCAGGTGAACTATCTATCGAAATATGATTTGATCTCAAGAAAGATGTCGTCGGGTAG	2300
QY	401	LysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGlu	420	QY	761	ArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeu	780
DB	1221	AAAGTTAAAGTGGACAAAGAAATACGCTGTACCAATGGAAATCTTGAAGTTTCATGAG	1280	DB	2301	CGTGCAGTGGCTCATGGACGGGAAGACAGATATTTTGGAGCAAAAGTTCTGTGAATTTG	2360
QY	421	LysProGlnArgTyrLysAenArgIleAspLeuValMetGlnAspLysPheLeuLysArg	440	QY	781	LeuArgGluPheAlaGluAenAenAspAenArgAlaProAlaHisIleValValTyrArg	800
DB	1281	AAAGCAAAAGATACAAAGATCGAATTTGATCTGGTATCGAAGCAAGTTTCTAAAGCGA	1340	DB	2361	CTCAGAGAAATTCGAGAAAAACAAACACAATTCGAGCACCGACGATATGTTAGTCTATCGA	2420
QY	441	AlaThrArgLysProHisAspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAsp	460	QY	801	AspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLys	820
DB	1341	GCTACACGAAAACTCACGACTACAAAGAAAAATACCCTTAAAAATGCTGAAAGAAATTTGAT	1400	DB	2421	GACGAGTTAGCGATTCGAGATGCTACGTGTGTAGTCATGATGAGCTTCGATCTTTTAAAA	2480
QY	461	PheSerSerGluGluLeuAenPheValGluArgPheGlyLeuCysSerLysLeuGlnMet	480	QY	821	SerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThr	840
DB	1401	TTCTCTTCTGAAGAGCTAAATTTTGTGAAGATTTTGGATTTATGCTCCAAATCTCAGATG	1460	DB	2481	AGCGAAGTAAACAAATTCATGTCGGAACGGGATCGAAGATCCAGAGCCGAAGTACACG	2540
QY	481	IleGluCysProGlyLysValLeuLysGluProMetLeuValAenSerValAenGluGln	500	QY	841	PheIleValIleGlnLysArgHisAenThrArgLeuLeuArgArgMetGluLysAspLys	860
DB	1461	ATCGAATGTCAGGAAAGGTTTTGAAAGACCAATGCTTGTGAATAGTGTAAATGAACAA	1520	DB	2541	TTCAATTTGATTCAGAAAAAGCACAAATACACGATTTGCTTCGAAAGATTGAAAAAGATAAG	2600
QY	501	IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAenValValProGlu	520	QY	861	ProValValAenLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys	880
DB	1521	ATTAATAATGACACCGAGTATCTGTGATTTCAAGAAAAACAAATTTGAATGTGTTCCCGAA	1580	DB	2601	CAAGTGGTCAATTAAGATCTTACTCTCTGCTGAAACAGATGTCGCTGCTGCTGTATAA	2660
QY				QY	881	GlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleValAenProSerSerGly	900


```
Db 2661 CAATGGGAGGATATGAAAGAAAGCAAGAACTGGAATGTGAACCCATCATCCGGA 2720
Qy 901 ThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHis 920
Db 2721 ACAACTGTGGATAAACTATCGTTTCGAATATACAAATTCGATTTTCTTGGCATCTCAT 2780
Qy 921 HisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGly 940
Db 2781 CATGGTCTCTTGTGTATCATCTCGTCCAGGACATTTACACTGTGTATGTATGACGATAAAGGA 2840
Qy 941 MetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCys 960
Db 2841 ATGAGCCAAAGTGAAGTCTATAAATGACCTACGAGCTTGCCTTCTCTGTAGATGT 2900
Qy 961 ArgLysProLysSerLeuProValHisTyrLysLeuAlaHisLeuSerCysGluLysAla 980
Db 2901 CGAAACCCATCTCGTTGCGCTGTTCCGGTTCATTATGCTCATTTATCATGTGAAAGCG 2960
Qy 981 LysGluLeuTyrArgThrTyrLysGluHisTyrLysLeuAlaHisLeuSerCysGluLysAla 1000
Db 2961 AAGAGCTTTATCGAATCTTCAAGGAACATTTACATCGGTGACTATGCACAGCCACGACT 3020
Qy 1001 ArgHisGluMetGluHisPheLeuGlnThrAsnValLysTyrProGlyMetSerPheAla 1020
Db 3021 CGACACGAAATGGAACTTTTCTCCAACTAACTGAAGTACCCCTGGAATGTGCTCGCA 3080
```

RESULT 3

```
US-11-144-985-2
; Sequence 2, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: INTERFERENCE
; CURRENT APPLICATION NUMBER: US/11/144,985
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(3080)
US-11-144-985-2
```

```
Alignment Scores:
Pred. No.: 0 Length: 3227
Score: 5349.00 Matches: 1020
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 16 Gaps: 0
```

US-10-645-746-3 (1-1020) x US-11-144-985-2 (1-3227)

```
Qy 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
Db 21 ATGTCTCGAATTTTCCGAAATGGAAAAAGGATTTATCGTCATCTCTCGATCCGGAG 80
Qy 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysLysVal 40
```

```
Db 81 ATGAAATGGCTTGCAGGCCCACTGGTAAATTCGACGCGCAAAATCTATAGAGAAAGTA 140
Qy 41 LeuLeuLeuValLeuLeuTyrPheLysPheSerSerLysLysLeuTyrAspArgGluTyrTyrCdu 60
Db 141 CTTCCTTTGGTAAATTTGGTCTTCAAGTTCTCCAGCAAAATTTTACGATCGGGAATACTACGAG 200
Qy 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
Db 201 TATGNACTGAAATGACAAAGGAAGTATTGAATAGAAAACCCAGGAAACCTTTCCCAAAA 260
Qy 81 LysThrGluLeuProLysProLysLeuAlaLysLeuPheThrGlnHisLeuArgHisGlu 100
Db 261 AAGACAGAAATTCGAATTCGCGATCGTCAAAATCTTCTGGCAACATCTTTCGCGCATCG 320
Qy 101 LysLysGlnThrAspPheLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
Db 321 AAGAGCGAGACAGATTTTATCTCGAAGACTATGTTTTGATGAAAAGGACACTGTTTAT 380
Qy 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValLys 140
Db 381 AGTGTGTTGTCGACTGAACACTGTACATCAAAATGCTGGTTTCGGAGAAAGTAGTAAA 440
Qy 141 LysAspSerGluLysLysAspGluLysAspLeuLysLysLysLysLysLysLysLysLys 160
Db 441 AAGGATTCGGAGAAAAAGATGAAAAGGATTTGGAGAAAAAAATCTTATACACAATGATA 500
Qy 161 LeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAspGlu 180
Db 501 CTTCACCTATCGTAAAAAATTTTACCTGAATTTAGTCGAGAAAAATCCGAAAAAAGAGCAA 560
Qy 181 GluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysValArgTyrAla 200
Db 561 GAAGCGAATCGGAGTTTACAAATTCCTGAGAAATGTTATGACCCAGAAAGTTGCTACGCG 620
Qy 201 ProPheValAsnGluGluLysValGlnPheAlaLysAsnPheValTyrAspAsnAsn 220
Db 621 CCTTTTGTGAACGAGAGATTAAGATACAAATTCGCGAAAAAATTTTGTGTACGATAAAT 680
Qy 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeuGlu 240
Db 681 TCANATTCGCGAGTTCTGATCGTTTTCAGATCCAAACAGATTCGAACAAATCAATTAGAA 740
Qy 241 ValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyLysGluLeuPheAsp 260
Db 741 GTAGCACCAAGAAATCGAAGCATGTTTGGAAATTTACATTCGAATCAAGAAATTTGTTTCGAT 800
Qy 261 GlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMet 280
Db 801 GGTGAACCTGTGCTCAATTTTGCATTTGTGCAATTAATCTTCTACCAATGCACGAAATG 860
Qy 281 SerLeuLeuAspTyrLysLeuLeuValLeuAspProGlnSerCysAsnAspAspValArg 300
Db 861 TCTCTTCTGGATTAUCTTCTCTTAATGTGACCCCGCTGCTGCTGATGATGATACGA 920
Qy 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaIleArgPro 320
Db 921 AAAGATCTTAAACAAAACTGATGCGCGGAAAAATGACAAATCAGACAGCCGCGCGCCA 980
Qy 321 ArgIleArgGlnLeuLeuGluAsnLeuLysLysCysAlaGluValTrpAspAsnGlu 340
Db 981 AGAATTCGACAAATTTATGGAAAAATTTGAAAGCTGAAATGCGCAGAGTTTGGGATAACGAA 1040
Qy 341 MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeu 360
Db 1041 ATGTGAGATTTGACAGACGACATCTGACATTTCTAGATTTTGTGCGAGGAAACTCTCTT 1100
Qy 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThr 380
Db 1101 GTTTATAAAGTCACTGGTAAATCGGACAGAGGAAATGCAAAAAAAGTACGATCTACTACA 1160
Qy 381 LeuPheLysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuVal 400
```

Db 1161 TTGTTCAAAATCTATGAGGAAAAAAGAGTTTCATTGAGTTTCCACCTACCTAGTC 1220
 Qy 401 LysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuValHisGlu 420
 Db 1221 AAAGTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATCGAAATCTTGAAGTTTCATGAG 1280
 Qy 421 LysProGlnArgTyrLysAenArgIleAspLeuValMetGlnAspLysPheLeuLysArg 440
 Db 1281 AGCCCAAAAGATACAAAGATCGAATTGATCTGGTGATCGAAGACAAAGTTCTTAAGCGA 1340
 Qy 441 AlaThrArgLysProHisAspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAsp 460
 Db 1341 GCTACACGAAACCTCACGACTACAAAGAAATACCTAAAAATGCTGAAAGAAATGGAT 1400
 Qy 461 PheSerSerGluGluLeuAenPheValGluArgPheGlyLeuCysSerLysLeuGlnMet 480
 Db 1401 TTCTCTTCTGAAGAGCTAAATTTTGTGAAGATTTGGATTTATGCTCCAAATCTCAGATG 1460
 Qy 481 IleGluCysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGln 500
 Db 1461 ATCGAATGTCAGGAAGAGTTTGAAGAGCCATGCTTGTGATGTGAATGAACAA 1520
 Qy 501 IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGlu 520
 Db 1521 ATTAATAATGACACCACTGATTCGTGATTTCAAGAAAAACAATTTGAATGTGTTCCGAA 1580
 Qy 521 LysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGlu 540
 Db 1581 AAAGAACTTTGTGTGCTGTTTTGTAGTCAACGAAACAGCGGAAATCATGCTTAGAA 1640
 Qy 541 GluAenAspValValLysPheTyrThrGluLeuIleGlyCysLysPheArgGlyIle 560
 Db 1641 GAGACGAGCTGTTAAGTTCTACCCGAACCTAATTTGGTGGTGCNAGTTCCGTGGAATA 1700
 Qy 561 ArgIleGlyAlaAsnGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAen 580
 Db 1701 CGAATTTGGTGCCAAATGAAACAGAGAGCGCAATCTATTATGTACGACGCGCAAAAT 1760
 Qy 581 GluTyrAlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAla 600
 Db 1761 GAATATGCTTCTACAAAAATTTGTACATAAATACCGGAATCGGTAGATTGAAATAGCC 1820
 Qy 601 AlaThrGluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMet 620
 Db 1821 GCACAGAGCGAGAAATATGTTGACGCTTCCGATTAAGACAAACAAAGCTTTAATG 1880
 Qy 621 PheIleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAsp 640
 Db 1881 TTCAATATCATTTCCAAACGACAACTGAATGCTTACGGTTTTGTGAAACATATTATTGCGAT 1940
 Qy 641 HisThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAla 660
 Db 1941 CACACCATCGGTGTAGTAAATCAGCAATATTTCTTGAACACAGTCAAAAGCTTTGGCA 2000
 Qy 661 SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAen 680
 Db 2001 TCATTAAGCAGCAGGAAGAGTCAAAACGAAATTTCTATCAATTTGCAATTTGAATCAAC 2060
 Qy 681 AlaLysLeuGlyGlyIleAenGlnLeuAenArgTyrPheGluIleAlaGluLysSerPro 700
 Db 2061 GCGAATTAAGAGGTATTAAACAGAGAGCTTGACTGGTTCAGAAATTCAGAAATATCACCA 2120
 Qy 701 GluGluLysGluArgGlyThrMetProLeuThrMetTyrValGlyIleAspValThr 720
 Db 2121 GAAGAAAAAGAGAGCGAAACAAATGCCATTAATCTATGTATGTGGAAATGATGTAAT 2180
 Qy 721 HisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlaSerIleAen 740
 Db 2181 CATCCAACTCTACAGTGGATTTGATTTCTATAGCGGCTGTAGTACGAGTATCAAT 2240
 Qy 741 ProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluCysArgProGlyGlu 760
 Db 2241 CCAGGTGGAATCTATTCGAAATATGATTTGACTCAAGAGAAATGTGTCGCCGTGAG 2300

Qy 761 ArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeu 780
 Db 2301 COTGARGTGGCTCATGGACGGGAAAGACAGATATTTTGGAAAGCAAGTTCTGTGAATTTG 2360
 Qy 781 LeuArgGluPheAlaGluAenAsnAspAenArgAlaProAlaHisIleValValTyrArg 800
 Db 2361 CTCAGAGAAATTCGACAGAAAAACAACGACATCGAGCAGCCAGCATATTGTAGTCTATCGA 2420
 Qy 801 AspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuAenArgSerLeuLys 820
 Db 2421 GACGAGTTAGCGATTCGAGATGCTGCTGTAGTTCATGATGAGCTTCGATCTTTAAAAA 2480
 Qy 821 SerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThr 840
 Db 2481 AGCGAAGTAAACAATTCATGTCGGAACGGGATCGAAGATCCAGAGCCGGAAGTACAGC 2540
 Qy 841 PheIleValIleGlnLysArgHisAenThrArgLeuLeuArgMetGluLysAspLys 860
 Db 2541 TTCATTGTGATTCAGAAAAAGACACAATACAGATTGCTTCGAAAGAAATGCAAAAGATAAG 2600
 Qy 861 ProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys 880
 Db 2601 CCAGTGGTCAATAAAGATCTTACTCTCTGCTGAACAGATGTCGCTGTTCTGCTGTATAA 2660
 Qy 881 GlnTyrGluGluAspMetLysGluSerLysGluThrGlyIleValAsnProSerSerGly 900
 Db 2661 CNAATGGAGGAGATATGAAAGAAAGCAAGAAATGGAATTTGTGAACCCATCATCCGA 2720
 Qy 901 ThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHis 920
 Db 2721 ACAACTGTGGATAAACTTATCGTTTCGAAATACAAATTCGATTTTCTTGGCATCTCAT 2780
 Qy 921 HisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGly 940
 Db 2781 CATGTTGCTTGTGGTACATCTGCTCCAGGACATTAACCTGTTATGTATGATACGATAAAGA 2840
 Qy 941 MetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCys 960
 Db 2841 ATGACCAAGATGAAGTCTTAATAATGACCTACGAGCTTCTTCTCTCTAGTAGTCT 2900
 Qy 961 ArgLysProIleSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAla 980
 Db 2901 CGAAAAACCATCTCGTTCGCTTTCGCTTCAATATGCTCATTTATCATGTGAAAAAGCG 2960
 Qy 981 LysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThr 1000
 Db 2961 AAAGAGCTTTATCGAACTTACAAAGAACATTAACATCGGTGACTATGCACAGCCAGGACT 3020
 Qy 1001 ArgHisGluMetGluHisPheLeuGlnThrAenValLysTyrProGlyMetSerPheAla 1020
 Db 3021 CGACACGAATGGAACATTTCTCCAACTAAGTGAAGTACCTCGTGAATGTCGTTCCGA 3080

RESULT 4

US-10-645-746-1
 ; Sequence 1, Application US/10645746
 ; Publication No. US20040265839A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mello, Craig C.
 ; APPLICANT: Tabara, Hiroaki
 ; APPLICANT: Grishok, Alla
 ; APPLICANT: Fire, Andrew
 ; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
 ; TITLE OF INVENTION: INTERFERENCE
 ; FILE REFERENCE: UMY-052DVI
 ; CURRENT APPLICATION NUMBER: US/10/645,746
 ; CURRENT FILING DATE: 2003-08-20
 ; PRIOR APPLICATION NUMBER: US 09/689,992
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: US 60/193,218
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: US 60/159,776
 ; PRIOR FILING DATE: 1999-10-15


```
QY 544 alVal-----LysPheTyrThrGlu 550
|||||
Db 1939 TTGTGTAAAGTGTCTTCTACGTAGATTATTCGGAATAATTTTCAGTAAAGTTCTACACCGAA 1998
QY 551 LeuileGlyCysLysPheArgGlyileAtrGileGlyAlaasnGluAenArgGlyAla 570
1999 CTAATTTGGTGGTTCGAAGTTCGTCGAATACGAATTTGGTCCCAATGAAACACAGAGAGCG 2058
QY 571 GlnSerileMetTyrAspAlaThrLysAsnGlu----- 581
2059 CAATCTATTATGTACACGCGACGGAATAATGAATATGCCGTAAAGTTTCAGAAATAATGAAA 2118
QY 582 -----TyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyileGly 595
2119 GTTTTTAAATATCATATTACAG-TTCTACAAAAATTTGACATAAATACCGAATCGGT 2177
QY 596 ArgPheGluileAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGlu 615
2178 AGATTTGAAATAGCCGCAACAGAGCGAAGAAATATGTTTGAACGCTCTCCCGATAAAGAA 2237
QY 616 GlnLysValLeuMetPheilelleSerLysArgGlnLeuAsnAlaTyrGlyPheVal 635
2238 CAATAAGTCTTAATGTTCAATATCATTTCCAAACGCAACTGAATGCTTACCGGTTTGTG 2297
QY 636 LysHisTyrCysAspHisThrileGlyValAlaAsnGlnHisileThrSerGluThrVal 655
2298 AAACATTATTCGATCACACCATCGGTGTAGCTAATCAGCATATTACTTTCGAAACAGTC 2357
QY 656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgilePheTyrGlnile 675
2358 ACATAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAACGAATTTTCTATCAAT 2417
QY 676 AlaLeuLysileAsnAlaLysLeuGlyileAsnGlnGluLeuAspTrpSerGluile 695
2418 GCATTTGAAATCAACCGSAATTTAGGAGGTATTAACGAGAGCTTGACTGGTCAGAAAT 2477
QY 696 AlaGluileSerProGluLysGluArgArgLysThrMetProLeuThrMetTyrVal 715
2478 GCAGAAATATCACCAGAGAAAGAAAGACGGAANAACAAATGCCATTAACTATGTATGTT 2537
QY 716 GlyileAspValThrHisProTherSertYrSerGlyileAspTyrSerileAlaAlaVal 735
2538 GGAATTTGATGTAATCATCCACCTCTACAGTGGAAATGATTTATCTATAGCGGTGTA 2597
QY 736 ValAlaSerileAsnProGlyThrileTyrArgAsnMetileValThrGlnGluGlu 755
2598 GTAGCGAGTATCAATCCAGGTGGAATCTATCGAAATATGATTTGTGACTCAAGAAGAA 2657
QY 756 CysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspileLeuGluAla 775
2658 TGTCGTCGCGGTGAGGTGCGTGGCTCATGACGCGGAAGAACAGATATTTGGNAGCA 2717
QY 776 LysPheValLysLeuLeuArgGluPheAlaGlu----- 786
2718 AAGTTCGTGAAATGCTCAGAGAAATTCGCAAGAGTGTGCTTGGATTTTAAAGAT 2777
QY 787 -----AsnAsnAspAsnArgAlaProAlaHi 795
2778 CTCTCGGATTTTTTAATTTTTTTGTAAACTTTCAGAACCAACGAACTCGAGACCGCGCA 2837
QY 795 sileValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG1 815
2838 TATTGTAGTCTATCGAGCGGAGTTAGCGATTCGGAGATGCTACGTGTAGTCATGATGA 2897
QY 815 uleuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
2898 GCTTCGATCTTTAAAAAGCGAAGTAAACAATTCATGTCGGAACGCGGATGGAGAAATCC 2957
QY 835 oGluProLysTyrThrPheileValileGlnLysArgHisAsnThrArgLeuLeuArgAr 855
2958 AGAGCGGAAGTACACGTTTATTGTGATTAGAAAAAGACAAATACACGATTCCTCGAAG 3017
QY 855 gMetGluLysAspLysProValValaAsnLysAspLeuThrProAlaGluThrAspValAl 875
```

```
Db 3018 AATGAAAAAGATAAGCCAGTGGTCAATAAAGATCTTACTCTGCTGAAACAGATGTCG 3077
QY 875 aValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyileVa 895
3078 TGTTCTGCTGTGTAAACAATGGAGAGGATATGAAGAAGAAACAAAGAACTTGAATTGT 3137
QY 895 IasnProSerSerGlyThrThrValAspLysLeuileValSerLysValTyrLysPheAspPh 915
3138 GAACCATCATCCGGAACAACATGTGGATAAACTTATCGTTTCGAATACAAATTCGATTT 3197
QY 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
3198 TTTCTTGGCATCTCATCTGTTGTTCTTGGTACATCTGTCAGGACATTTACACTGTTAT 3257
QY 935 tTyrAspAspLysGlyMetSerGluAspGluValTyr----- 947
3258 GTATGACGATAAAGGAATGAGCAAGATCAAGTCTATGTAAAGCGTTTGAATAGCAGTTA 3317
QY 948 -----LysMetT 950
3318 GCGATTTTAGGATTTTGTAAATCCGCATATAGTTATTATAAAAAAATGTTTCAGAAAATGA 3377
QY 950 hrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProileSerLeuProValProV 970
3378 CCTACGGACTTGTCTTCTCTGCTAGATGTGGAACCCCATCTCGTTGCCCTGTTCCGG 3437
QY 970 alHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluH 990
3438 TTCATTATGCTCATTTATCATGTGAAAAAGCGAAAGAGCTTTATCGAACTTACAAGGAAC 3497
QY 990 isTyrileGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnT 1010
3498 ATTACATCGGTGACTATGCACAGCCAGACTCGACACGAAATGGAACATTTTCTCCAAA 3557
QY 1010 hrAsnValLysTyrProGlyMetSerPheAla 1020
3558 CTAACGTGAAGTACCCTGGAATGTCGTTGCGCA 3589
RESULT 5
US-10-645-735-1
; Sequence 1, Application US/10645735
; Publication No. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UM-G-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3719
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-645-735-1
Alignment Scores:
Pred. No.: 0 Length: 3719
Score: 5001.00 Matches: 1016
Percent Similarity: 85.2% Conservative: 1
Best Local Similarity: 85.1% Mismatches: 3
Query Match: 93.5% Indels: 177
```

DB:	10	Gaps:	10
US-10-645-746-3 (1-1020) x US-10-645-735-1 (1-3719)			
Qy	1	MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspPro---	19
Db	21	ATGTCCTCGAATTTTCCGAATTCGGAATTCGGAATTCGGAATTCGTCATCTCTCGATCCGGTA	80
Qy	20	-----GluMetLys	22
Db	81	TGATCAATATTAGCAGCTATAAGATATATAAGTTTGATATTAATATTATAGGAGATGA	140
Qy	22	sTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysLysValLeuLe	42
Db	141	ATGGCTTGGAGGCCCACTGGTAATCCGACGGCAAAATTCATGAGGAAGAAGTACTTCT	200
Qy	42	uLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrG1	62
Db	201	TTTGGTAAATTTGGTTCAAGTTCTCCAGCAAAATTTACGATCGGAATACCTACGAGTATGA	260
Qy	62	uValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysTh	82
Db	261	AGTGAATATGACAAAGGAAGTATTGAATAGAAACCAGGAACCTTTCCCAAAAAGAC	320
Qy	82	rGluIleProIle-----ProAs	88
Db	321	AGAAATTCCAATGTAAGTCTGTAAATTAGTCAAACTAATTTATTTTTCAGTCCCGA	380
Qy	88	pArgAlaLysPheTrpGlnHisLeuArgHisGluLysLysGlnThrAspPheIleLe	108
Db	381	TCGTGCAAAACTCTCTCGCAACATCTTCGGCATGAGAAGACGACAGATTTTATCT	440
Qy	108	uGluAspTyrValPheAspGluLysAspThrValTyrSerValCysArgLeuAsnThrVa	128
Db	441	CGAAGACTATGTTTGTGATGAAAGGACACTGTTTATAGTGTGTCGACTGAACACTGT	500
Qy	128	lThrSerLysMetLeuValSerGluLysValValLysLysAspSerGluLysLysAspG1	148
Db	501	CACATCAAAATGCTGTTTCGGAGAAAGTAGTAAAGAGATTCGGAGAAAGAGATGA	560
Qy	148	uLysAspLeuGluLysLysIleLeuTyrThrMetIleLeuThrTyrArgLysLysPheH1	168
Db	561	AAAGGATTTGGAGAAAAATCTTATACACATGATACTTACCTATCGTAAATAATTTCA	620
Qy	168	sLeuAsnPheSerArgGluAsnProGluLysAspGluGluLysValLysAsnArgSerTyrLysPh	188
Db	621	CCTGAACTTTAGTCGAGAAAAATCCGGAAGAACGAGGAGATTAAGTGTGAGTTGCAAT	680
Qy	188	eLeuLys-----AsnValMe	193
Db	681	CCTGAA-GGTTTATGAAAAACGCACTTATAACAAATAATTTAGCTTTTCAGAAATGTTAT	739
Qy	193	tThrGlnLysValArgTyrAlaProPheValAsnGluGluIleLysVal-----	209
Db	740	GACCCAGAAAGTTTCGTACCGCTTTTGTGAACGAGGAGATTAAGTGTGAGTTGCAAT	799
Qy	210	-----GlnPheAlaLysAsnPh	215
Db	800	AATAATAATAATAATCACCTCAACTCATTTATATATATTTTAAAGCAATTCGCGAAAAATTT	859
Qy	215	eValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPh	235
Db	860	TGTGTACGATAATAATTCATTTCTCGAGTTCCTGAATCGTTTTCAGATCCAAACAGATTT	919
Qy	235	eGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyI1	255
Db	920	CGAACATCATTAGAAGTAGCACCAAGATCGAAGCATGGTTTGGAAATTACATTGCAAT	979
Qy	255	eLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIle-----	269
Db	980	CRAAGAAATGTTTCGATGGTGAACCTGCTGCTCAATTTTGCAAG-TAAGTTTGAGAAACTGC	1038
Qy	270	-----ValAspLysLeuPheTyrAsnAlaProLys	279
Db	1039	GATAAAAAATCATGTGTGATTTTGTGAAGTTGTGATATAAATAATTTTACAAATGACCGAA	1098
Qy	279	sMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAspAspVa	299
Db	1099	AATGTCTCTCTCGATTAATCT	1158
Qy	299	lArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaLaAr	319
Db	1159	ACGAAAAGATCTTTAAAAACAAATCTGATGGCGGAAAAAATGACAAATCAGACAAACGCGCG	1218
Qy	319	gProArgIleArgGlnLeuLeuGluAsnLeuLysCysAlaGluValTrpAspAs	339
Db	1219	GCCAAGAAATTCGACAATTTATGGAAAAATTTGAAGCTGAAATGCGCAGAAATTTGGGATAA	1278
Qy	339	nGluMet-----SerArgL	344
Db	1279	CGAAATGTTAGTTTAAATTTTAAATTAATATATAAATTTGATTTTTCAGGTCGAGAT	1338
Qy	344	eThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysV	364
Db	1339	TGACAGAACGACATCTGACATTTCTAGATTTGTGCGAGGAAAACTCTCTTGTTTTATAAAG	1398
Qy	364	alThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysI	384
Db	1399	TCACTGGTAAATCGGACAGAGGAAGAAATGCAAAAAAGTACGATACTTACATTTGTTCAAAA	1458
Qy	384	lTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysS	404
Db	1459	TCTATGAGGAACAAAGTTCAATTTGAGTTTCCCACTACCACTAGTCAAGTTTAAAA	1518
Qy	404	erGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnA	424
Db	1519	GTGGAGCAAAAGAAATACGCTGTACCAATGGAACATCTTGAAGTTTCATGAGAAGCACAAA	1578
Qy	424	xGlyLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgL	444
Db	1579	GATCAAGAAATCGAATTTGATCTGGTGTGATGCAAGCAAGTTTCTAAAGCGAGCTACACGA	1638
Qy	444	ysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLysAspPheSerSerG	464
Db	1639	AACTTCAAGCTACAAAGAAATACCTTAAATGCTGAAAGAAATTTGATTTCTCTCTG	1698
Qy	464	luGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysP	484
Db	1699	AAGAGCTAAATTTTGTGTAAGATTTGATTTGCTCCAACTTCAGATGATCGAATGTC	1758
Qy	484	roGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetT	504
Db	1759	CAGGAAAGGTTTGAAGAGCCCAATGCTGTGAATAGTGTAAATGAACAAATTTAAATGA	1818
Qy	504	hrProValIleArgGlyPheGlnLysGlnLeuAsnValValProGluLysGluLeuL	524
Db	1819	CACCACTGATTCGTGATTTCAAGAAAAACAATTTGAATGTGGTTCCCGAAAAAGACTTT	1878
Qy	524	ysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAspV	544
Db	1879	GCTGTGCTGTTTGTAGTCAACGAAACAGCGGAAAAATCCATGCTTTAGAGAGAACGACG	1938
Qy	544	alVal-----LysPheTyrThrGlu	550
Db	1939	TTGTGTAAAGTGTTCCTAGTAGATTTTCCGAAAAATTTTTCAGTAGTAAATTTCTACACCGAA	1998
Qy	551	LeuIleGlyGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyVala	570
Db	1999	CTAATTTGGTGTTCGAATTCGTGGAATACGAATTTGGTGGCAATGAAACAGAGAGCG	2058
Qy	571	GlnSerIleMetTyrAspAlaThrLysAsnGlu-----	581
Db	2059	CAATCTATTATGTACGACGCGCAAAAAATGAATATGCGGTAAATTTTCAGAAAAATGAAA	2118
Qy	582	-----TyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGly	595

```
Db 2119 GTTTTTAAATATCATATTATACAG-TTCTACAAAAATTGTACACTAAATACCGAATCGGT 2177
Qy 596 ArgPheGluIleAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGlu 615
Db 2178 AGATTTTGAATAGCCGCAACAGAGCGAAGAAATATGTTTGAACGCTCTCCCGATAAAGAA 2237
Qy 616 GlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaThrGlyPheVal 635
Db 2238 CAAAAGTCTTAATGTTTCAATATATCTTCCAAACGCACTGAATGCTTACCGGTTTGTG 2297
Qy 636 LysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVal 655
Db 2298 AAACATTATTGGCATCACACCATCGGTGTAGCTAATCAGCATATTACTTCTGAAACAGTC 2357
Qy 656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIle 675
Db 2358 ACAAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAACGAATTTTCTATCAAAAT 2417
Qy 676 AlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIle 695
Db 2418 GCATTGAAATCAACCGGAAATTAGGAGGTATTAAACGAGAGCTTGACTGGTCAGAAAT 2477
Qy 696 AlaGluIleSerProGluLysGluArgArgLysThrMetProLeuThrMetTyrVal 715
Db 2478 GCAGAAATATCACCAGAGAAAGAAAGACGCGAACAATGCCATTAACTATGTATGTT 2537
Qy 716 GlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaVal 735
Db 2538 GGAATTTGATGTAACATCCACCTCCTACAGTGGAAATTTGATTTCTATAGCGGCTGTA 2597
Qy 736 ValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGlu 755
Db 2598 GTAGCGAGTATCAATCCAGTGGAGTATCTATCGAATATGATTTGTGACTCAAGAAGAA 2657
Qy 756 CysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAla 775
Db 2658 TGTGTCGCCGTCGAGCGTCAGTGGCTCATGGACGGGAAAGAACAGATATTTTGGAAAGCA 2717
Qy 776 LysPheValLysLeuLeuArgGluPheAlaGlu----- 786
Db 2718 AAGTTGCGTGAATTTGCTCAGAGAAATTCGAGAAAGTGAGTTGCTTGGAGTATTTAAAGAT 2777
Qy 787 -----AsnAsnAspAsnAtqAlaProAlaHi 795
Db 2778 CTCTGGGATTTTAAATTTTTTGTAAACCTTCAGAAACAACGACATCGAGACCGCGCA 2837
Qy 795 sileValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGl 815
Db 2838 TATTGTAGTCTATCGAGACGGAGTTAGCGATTCCGGAGATGCTACGTGTTAGTCATGATGA 2897
Qy 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
Db 2898 GCTTCGATCTTTAAAGCGAAGTAAACAATTCATGTCGGAACCGGATGGAGAAAGATCC 2957
Qy 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgAr 855
Db 2958 AGAGCCCAAGTACACGTTTCAATGTGATTCAGAAAAGACACAATACAGATTCCTTCGAAG 3017
Qy 855 gMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAl 875
Db 3018 AATGGAAAAAGATAAGCCAGTGGTCAATAAAGATCTTACTCTGCTGAAACAGATGTCG 3077
Qy 875 aValAlaLalValLysGlnTrpGluLysPheMetLysGluSerLysGluThrGlyIleVal 895
Db 3078 TGTGTGCTGTTGTAAACAATGGAGAGGATATGAAAGAAAGCAAAAGAACTCGAATGCT 3137
Qy 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
Db 3138 GAACCATCATCCGGAACAACATGTGGATAACTTATCGTTTCGAAATACAAATTCGATTT 3197
Qy 915 ePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
Db 3198 TTTCTTGGCATCTCATCGTGGTCTCTGTGATCACTCTCGTCCAGGACATTACACTGTTAT 3257

RESULT 6
US-11-144-985-1
; Sequence 1, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UMY-052CN
; CURRENT APPLICATION NUMBER: US/11/144,985
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3719
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-11-144-985-1

Alignment Scores:
Pred. No.: 0 Length: 3719
Score: 5001.00 Matches: 1016
Percent Similarity: 85.2% Conservative: 1
Best Local Similarity: 85.1% Mismatches: 3
Query Match: 93.5% Indels: 177
DB: 16 Gaps: 10

US-10-645-746-3 (1-1020) x US-11-144-985-1 (1-3719)
Qy 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspPro--- 19
Db 21 ATGTCTCGAATTTTCCGAAATTTGAAAGAAAGGATTTATCGTCATTTCTCTCGATCCGGA 80
Qy 20 -----GluMetLy 22
Db 81 TGATCAATTTATAGCAGCTATAGATATATAGTTTGATTAATATATATAGAGATGAA 140
Qy 22 eTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysValLeuLe 42
```



```
QY 656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIle 675
Db 2358 ACAAAAGCTTTGGCATCTACTAAGGCACGAGAAAGGATCAAAACGAATTTTCTATCAAAAT 2417
QY 676 AlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTyrSerGluIle 695
Db 2418 GCATTGAANAATCAACCGGAATTAAGGAGTATTAAACCGAGCTTGACTGGTCAGANAAT 2477
QY 696 AlaGluIleSerProGluLysGluArgArgLysThrMetProLeuThrMetTyrVal 715
Db 2478 GCAGAAATATCACAGAAAGAAAGAGACGCGAAACAATGCCATTAACTATGTATGTT 2537
QY 716 GlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaVal 735
Db 2538 GGAATTGATGAACATCATCAACCTCTACAGTGGAAATTTGATTATTCTATACGGCTGTA 2597
QY 736 ValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGlu 755
Db 2598 GTAGCGAGTATCAATCCAGGTGGACTATCTATCGAATATGATTGTGACTCAAGAGAA 2657
QY 756 CysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAla 775
Db 2658 TGTCTGCCGCTGAGCGTGCAGTGGCTCATGGACGCGAAAGAACAGATATTTTGGAGCA 2717
QY 776 LysPheValLysLeuLeuArgGluPheAlaGlu----- 786
Db 2718 AAGTTCGTGAAATGCTCAGAGAAATTCGCAGAAAGTGAGTTGCTTGGAGTATTTAAAGAT 2777
QY 787 -----AenAsnAspAsnArgAlaProAlaHis 795
Db 2778 CTCTGGATTTTAAATTTTTTTGTAACCTTCAGAACACAGCAATCGACACAGCGCA 2837
QY 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGlu 815
Db 2838 TATTGTAGTCTATCGACAGCGAGTTAGCGATTTCGGAGATGCTACGTTGTAGTTCATGATGA 2897
QY 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
Db 2898 GCTTCGATCTTTAAAGCGAAGTAGTAACAATTCATGTCGGAACGGGATGGAGAGATCC 2957
QY 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgAr 855
Db 2958 AGAGCGCAAGTAGACGTTCACTTGTGATTCAGAAAGACACATACACGATTCCTCGAAG 3017
QY 855 gMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAl 875
Db 3018 AATGGAAGAAAGATAAGCCAGTGGTCAATTAAGATCTTACTCTGCTGCTGAAACAGATGTCG 3077
QY 875 aValAlaAlaValLysGlnTrpGluLysPheMetLysGluSerLysGluThrGlyIleVa 895
Db 3078 TGTTCGCTGCTGTTTAAACAATGGAGAGGATATGAAGAAAGCAAGAACTGGAAATGT 3137
QY 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
Db 3138 GAACCATCATCCGGAACAACGTGGATAAATCTATCGTTTCGAAATACAAATTCGATTT 3197
QY 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
Db 3198 TTTCTTGGCATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3257
QY 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyr----- 947
Db 3258 GTATGACGATAAAGGAATGAGCAAGATGAAGTCTATGTAAGCGTTTTCGAAATAGCAGTGA 3317
QY 948 -----LysMet 950
Db 3318 GCGATTTTAGGATTTTGTAAATCGCATATAGTTATTATTAATAAATAATGTTTCAGAAATGA 3377
QY 950 hTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProv 970
Db 3378 CCTACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3437
QY 970 aHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluH 990
```

```
Db 3438 TTCATTATGCTATTATATCATGTGTAAGAAAGCGAAAGAGCTTTATCGACTTACAAAGAAC 3497
QY 990 isTyrIleGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnT 1010
Db 3498 ATTACATCGGTGACTATGCACAGCCACGAGCTCGACACGAAATGGAACATTTTCTCCAAA 3557
QY 1010 hrAsnValLysTyrProGlyMetSerPheAla 1020
Db 3558 CTAACGTGAAGTACCCTCGAATGTCGTTGCGCA 3589

RESULT 7
US-10-007-078-3
; Sequence 3, Application US/10007078
; Publication No. US20030105042A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF EIF2C1 EXPRESSION
; FILE REFERENCE: RTS-0236
; CURRENT APPLICATION NUMBER: US/10/007,078
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)...(2787)
US-10-007-078-3

Alignment Scores:
Pred. No.: 1,13e-41 Length: 7478
Score: 577.00 Matches: 237
Percent Similarity: 40.2% Conservative: 173
Best Local Similarity: 23.2% Mismatches: 370
Query Match: 10.8% Indels: 243
Db: 6 Gaps: 38

US-10-645-746-3 (1-1020) x US-10-007-078-3 (1-7478)
QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu----- 37
Db 209 ATGGGATGAA- GCGGACCTTCGGGACGACGCTCGGGCGCTTACTGCCCCCTGCGAG 267
QY 38 -----LysLysValLeuLeu 42
Db 268 CAGGTGTTCCAGGCACCTCGCGGCTGGCATTGGCATTGGGGAACCAATCAAGCTC 327
QY 43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGlu 62
Db 328 CTGGCCAATTACTTTGAGGTGGACATCCCTAAGATCGACGCTGTACCACTACGAGGTGGAC 387
QY 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysPheProPheProLysLysThr 82
Db 388 ATCAAGCCGATAG- -----TGTCCCCGTAGATGC 417
QY 83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
Db 418 AACCGGAAGTGGTGAA- -----TACATGCTCAGCATTTCAAGCCT 459
QY 103 GlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLysAspThrValTyr 120
Db 460 CAG-----ATCTTTGGTGATCGCAAGCCTGTGTATGATGGAAGAAAGAACATTTAC 510
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal----- 138
Db 511 -----ACTGTACAGCATGCCCCATGGCAACGAGGTTCGACTTT 552
QY 139 ---ValLysLysAspSerGluLysLysAspGluLysLysAspLeuGluLysLysIleLeuTyr 157
Db 553 GAGGTGACAATCCCTCGGGGAAGGAAGGAT---CGAATCTTTAAGGTCTCTCAAGTGG 609
```


158 ThrMetIleLeuThrThrArgLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
160 CTAGCCATTGTGAGCTGGCGAATCTGCTGAGGCCCTGTGTGTCAGCGCCAGATCCCTGTT 669
178 LysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysVal 197
179 CCCTTGAGTCTGTGCAAGCCTGGATGTGGCCATGAGGACCTGGGATCC-----ATG 723
198 ArgTyrAlaProPheValAsnGluGluLysValGlnPheAlaLysAsnPheValTyr 217
200 AGGTACACCCCTGTG----- 738
218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
219 ---GGCGCTCTCTTCTCAGCGCTGAGGCTTACTACACCCG----- 780
238 SerLeuGluValAlaProArgIleGluAlaTyrPheGlyIleTyrIleGlyIleLysGlu 257
240 ---CTGGGGGTGGCGGCAAGTCTGGTTCGGCTTTCACAGTCTGTGCGCCT 831
258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
260 ---TGGAAGATGATCTCAACATTGATGCTCTCAGCCACTGCTTTTATAAGCA 888
278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuValAspProGlnSerCysAsnAsp 297
280 ---CAGCCAGTGTGATTTGATGTTGATGCTGAGGTGCTGAGCATCAGGAACATAGATGAG 942
298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
300 ---CAGCCCAAG-----CCCTCAGGAGCTCT 966
318 AlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys-----CysAla 334
320 ---CAGCGCTTCTTCAACAGGAGATCAAGGCTGAGGTGGAAGTCAACCCACTGTGA 1026
335 GluValTyrAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr 350
351 PheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArg 370
371 GlyArgAsnAlaLysLysTyrAspThrLeuPheLysIleTyrGluGluAsnLysLys 390
391 Phe---IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyr 409
392 CTTTCAGCTCAAGTATCCCATCTGCCCTGCTCAAGTTGGCCAGGAACAAAGCATACC 1218
410 AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIle 429
429 TACCTTCCCTCAGAGTCTTACATT---GTGGCTGGGAGGCTGTATTAAGAAGCTG 1275
430 AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys 449
449 ACCGACACACAGACCTCGACCATGATAAAGCCACAGCTAGATCCGCTCCAGACAGACAG 1335
450 GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal 469
469 GAGGAGATCAGTCCGCTGATGAAGAATGCCAGTACCACTTAGATCCC-----TACATC 1389
470 GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
489 CAGGAATTTGGGATCAAGTGAAGATGATCATCGGAGGTGACAGGGCGAGTGTGCGCG 1449
490 GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro----- 505
490 GCGCCCATCTTGCAGTACGCGCGCGGACCGGCCCATTTGCCACACCCCATCAGGGTGTCT 1509

506 ---ValIleArgGlyPheGlnGluLysGlnLeu---AsnValValProGluLysGluLeu 523
510 TGGGACATGGGGGG-----AAACAGTTCTTACAAATGGATGATCAAAAGTCTGG 1560
524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
544 GCATCGCTTGTGCGACCCCAAAA---CAGTGTGCGAAGAG----- 1602
544 ValValLysPheTyrThrGluLeuIle---Gly-----GlyCysLysPhe 557
557 GTGCTCAAGAACTTCAACAGACCGCTGCGGAAGATTTTCAAGGATCGGGGATCGCTATC 1662
557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
577 CAGGGTCAACCTTTGTTTCTGCAAAATATGACAGGGGCA---GACACGCTG----- 1710
577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPhe 597
597 ----- 1710
597 eGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLys 617
617 ---GAGCCTATGTTCCGGCATCTCAAGAACACCTACTCAGG 1748
617 sValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis 637
637 GCTGCAGCTCATTTATGTCATCTGCCAGGAAGACCGCGGTGTATGCTGAGGTGAACG 1808
637 sTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
657 TGTCGGAGATACACTCTTGGGAATGCTACGAGTGTGTGAGTGAAGAACGTGGTCAA 1868
657 sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLe 677
677 ---GACC-----TCACCTCAGACTGTGTCACACCTCTGCCT 1901
677 LysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTyrSerGluIleAlaGln 697
697 CAAGATCAATGTCAAACTTGTGTGCATTAACAATCCTA----- 1941
697 uIleSerProGluGluLysGluArgLysThrMetProLeuThrMetTyrValGlyIle 717
717 ---GTCCACACAGCGCTCTCGCGTTTTCACAGCCAGCTG---ATATTCTGGGAGC 1994
717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValAlaVal 737
737 AGATGTTACACACCCCGCAGCGGATGGGAAAAAACCTTCTATCACAGCAGTGTAGG 2054
737 aSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGluCysArg 757
757 CAGTATGATGTCACCCCGCAGCGATCTGTGTACTGTGCGGGTACAG----- 2103
757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
775 ---CGACCGCGCAAGAGATCAATGAAGACTTGTGC----- 2135
775 aLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794
794 CTACATGGTGGTGAGCTCTCATCCAAATCTTCAAGTCCACCCCGTTTCAAG---CCTAC 2192
794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
814 CGCATCATCTTCTACCGAGATGGGTGCTGAGGCGCAGCTACCCCGAGTACTCCATTA 2252
814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834
834 TGAGCTACTGGCCATCTGCTGATGCTGCATCAACTG-----GAAAAGGA 2297
834 pProProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArg 854
854 CTACACGCTGGGATCATTTATTTGTTGGTGCAGAAACGCCATCACACCCGCTTTTC-- 2355
854 gArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspVa 874

```

Db      2356  -TGTGCTGACAGAAATGGGAGAGTGGTAACATCCAGCT------ 2403
QY      874  lalaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyI 894
Db      2403  ----- 2403
QY      894  eValAenProSerSerGlyThrThrValAspLysLeuIleValSerLysTyLysPheAs 914
Db      2404  -----GGACACACAGTGGACACCAACATCACCACCCATTGAGTTGA 2447
QY      914  pPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyThrVa 934
Db      2448  CTTCTATCTGTGAGCCAGCAGCATCCAGGCGACCGCCATCCCATTTACTATGT 2507
QY      934  lMetTyAspAspLysGlyMetSerGlnAspGluValTyLysMetThrTyGlyLeuAl 954
Db      2508  TCTTTGGGATGACACCGTTTCACAGCAGATGAGTCCAGATCCTGACGCTACGAGTGTG 2567
QY      954  aPheLeuSerAlaAtqCysArgLysProIleSerLeuProValProValHisTyAlaHi 974
Db      2568  CCACACTTACGTACGATGACACGCTCTCTCTATCCAGCACCCTGCTACTATGCCCG 2627
QY      974  sLeuSerCysGluLysAlaLysGluLeuTyArgThrTyLysGluHisTyIleGlyAs 994
Db      2628  CTGTGTGGCTTCCGGGACGAG---TACCACCTGGTGGACAGGACATGACAGTGGAGA 2684
QY      994  p 994
Db      2685  G 2685

```

RESULT 8

```

US-11-035-669-3
; Sequence 3, Application US/11035669
; Publication No. US20050182015A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF EIF2C1 EXPRESSION
; FILE REFERENCE: BNDL-001IUS.P1
; CURRENT APPLICATION NUMBER: US/11/035,669
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 09/793,807
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/953,611
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 09/954,679
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 10/007,078
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)... (2787)
US-11-035-669-3

Alignment Scores:
Pred. No.: 1,13e-41 Length: 7478
Score: 577.00 Matches: 237
Percent Similarity: 40.2% Conservative: 173
Best Local Similarity: 23.2% Mismatches: 370
Query Match: 10.8% Indels: 243
DB: 13 Gaps: 38

US-10-645-746-3 (1-1020) x US-11-035-669-3 (1-7478)

QY      21  MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrglu----- 37

```

```

Db      209  ATGGATGGAA-GCGGGACCCCTCGGAGCAGCTGCGGGCGCTTACTCTGCCCCCTCGCAG 267
QY      38  -----LysLysValLeuLeu 42
Db      268  CAGGTGTTCCAGGCACCTCGCGGCTGGCATTGGCAGCTGTGGGGAACCAATCAAGCTC 327
QY      43  LeuValAenTrpPheLysPheSerSerLysIleTyArgArgGluTyTrGluTyrglu 62
Db      328  CTGGCCATTTACTTTGAGGTGCGACATCCCTAAGATCGAGCTGTACCACCTACGAGGTG 387
QY      63  VallysMetThrLysGluValLeuAenArgLysProGlyLysProPheProLysLysThr 82
Db      388  ATCAAGCCGGATAAG-----TGTCCTCCCTAGATGC 417
QY      83  GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
Db      418  AACCGGGAAGTGGTGGAA-----TACATGGTCCAGCATTTCAAGCCT 459
QY      103  GlnThrAspPheIleLeuGluAspTy-----ValPheAspGluLysAspThrValTy 120
Db      460  CAG-----ATCTTTGGTGATCGCAAGCTGTGTATGATGGAAGAAGAACAATTAC 510
QY      121  SerValCysArgLeuAenThrValThrSerLysMetLeuValSerGluLysVal----- 138
Db      511  -----ACTGTACAGCAGCTGCCCATTTGGCAAGCAGCGGTGCGACTTT 552
QY      139  ---VallysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTy 157
Db      553  GAGGTGACAATCCCTCGGGGAAGGAGAT---CGAATCTTTAAGGTCTCCATCAAGTGG 609
QY      158  ThrMetIleLeuThrTyArgLysLysPheHisLeuAenPheSerArgGluAenProGlu 177
Db      610  CTAGCCATTGTGAGCTGGCGAATGTCGATGAGGCCCTGGTCAGCGGCCAGATCCCTGT 669
QY      178  LysAspGluGluAlaAenArgSerTyLysPheLeuLysAsnValMetThrGlnLysVal 197
Db      670  CCTTTGGAGTCTGTCAAAGCCCTGATGTGGCATGAGGACCATCGCATCC-----ATG 723
QY      198  ArgTyAlaProPheValAsnGluLysValGlnPheAlaLysAsnPheValTy 217
Db      724  AGGTACACCCCTGTG----- 738
QY      218  AspAenAenSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
Db      739  ---GGCGCTCTCTCTCTCCACCGCTGAGGCGCTACTACCACCG----- 780
QY      238  SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyIleGlyLysGlu 257
Db      781  -----CTGGGGGGTGGCGCGAAGTCTGGTTCTGGCTTTCCACGAGTCTGTGCGCCCT 831
QY      258  LeuPheAspGlyGluProValLeuAenPheAlaIleValAspLysLeuPheTyAenAla 277
Db      832  GCCATG---TGAAGATGATGCTCAACATTGATGTCTCAGCCACTGCCCTTTTATAGGCA 888
QY      278  ProLysMetSerLeuLeuAspTyTrLysLeuLeuIleValAspProGlnSerCysAenAsp 297
Db      889  -----CAGCCAGTGTGAGTTTCATGCTGAGGTGCTGGACATCAGGAACATAGATGAG 942
QY      298  AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
Db      943  CAGCCCAAG-----CCCTCAGGAGTCT 966
QY      318  AlaArgProArgIleArgGlnLeuLeuGluAenLeuLysLeuLys-----CysAla 334
Db      967  CAGCGCTTCGCTTCCCAAGAGATCAAGGGCCCTGAGGTGGAAGTCAACCCACTGTGA 1026
QY      335  GluValTrpAspAen-----GluMetSerArgLeuThrGluArgHisLeuThr 350
Db      1027  CAGATGAAGAGGAAGTACCGCGTGTGTAATGTTACCGCTCGCCCTGCTGCTAGCCATCAGACA 1086
QY      351  PheLeuAspLeuCysGluGluAenSerLeuValTyLysValThrGlyLysSerAspArg 370
Db      1087  TTCCCTTTACAGCTGGAGAGTGGACAGACTGTG----- 1119

```


; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Pilgrim, Marsha
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Reuber, Lynne
 ; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
 ; FILE REFERENCE: MBI-010
 ; CURRENT APPLICATION NUMBER: US/09/533,029
 ; CURRENT FILING DATE: 2000-03-22
 ; EARLIER APPLICATION NUMBER: 60/125,814
 ; EARLIER FILING DATE: 1999-03-23
 ; NUMBER OF SEQ ID NOS: 121
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 105
 ; LENGTH: 2910

; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; OTHER INFORMATION: G1149
 US-09-533-029-105

Alignment Scores:
 Pred. No.: 4e-42 Length: 2910
 Score: 575.50 Matches: 237
 Percent Similarity: 39.0% Conservative: 159
 Best Local Similarity: 23.3% Mismatches: 337
 Query Match: 10.8% Indels: 283
 DB: 3 Gaps: 41

US-10-645-746-3 (1-1020) x US-09-533-029-105 (1-2910)

```

; 895 GATTATGTCCTCTGTTGGGAGGTCTTTT-----TTCCACACTAGCTTTGGGA 939
; 240 -----GluValAlaProArgIleGluAlaTrpPheGlyIle 251
; 940 AAGGACGCAAGAGATGGTAGGGGTGAGCTTGGAGATGGTATTGAGTACTGGAGAGGTAT 999
; 252 TyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAsp 271
; 1000 TTCCAAAGTCTAAGG---CTGACTCAGATGGGTTGCTCTCTGAACATTGACGTTTCAGCA 1056
; 272 LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAsp 291
; 1057 AGATCATTTTAT-----GAA 1071
; 292 ProGlnSerCysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
; 1072 CCGATTGTTGTCACCTGACTTTATTAGCAAGTTTCTGAATATATAGGGACTTA-----AAC 1125
; 312 MetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuGluLeuAsnLeuLysLeu 331
; 1126 AGGCCACTTAGAGCTCAGATCAGCTTAAGGTGAAGAAAGTTTGTGAGGACACTGAAGATT 1185
; 332 LysCysAlaGluValTrpAspAsnGluMetSerArg----- 343
; 1186 AAGTTGCTTCAC---TGGAAACGGCACAAAAAGTGCACAAATTAGTGGGATTTCTAGTCTA 1242
; 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyr 362
; 1243 CCCATCAGGGAGCTAAGGTTCTCTGGAGGACAAATCAGAGAAAGCGGTGTGTTCAATAT 1302
; 363 LysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPhe 382
; 1303 -----TTTGTGAAAAATATTAAT----- 1320
; 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
; 1321 -----TATAGA-----GTGAAATACCAAGGCTCTACCTGCTATTCAACA 1359
; 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422
; 1360 GGGAGTGCACAAAGCCCGCTCTACCTACCAATGGAGCTCTGCCAAATTCACGAA---GGG 1416
; 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
; 1417 CAAGATATACCAAAAGGCTCAATGAGAGCAAGTGCATGTCATGCTATAAAGTACTACTGC 1476
; 443 ArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSer 462
; 1477 CAACGACCCCTGTAT---AGAGAGAACTCGATCAAAAACCTTGTTGTGAAAAATTAATTAC 1533
; 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu 482
; 1534 AATGATGATCTGAGC-----AAGGAGTTTGGGATGTCTAGTCACTACCAACTAGCCTCG 1587
; 483 CysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLys 502
; 1588 ATTGAAGCTCGTGTACTTCCCGCCCGATGTTGAAGTACCATGACAGTGTGAAGAGAAA 1647
; 503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGlu 522
; 1648 ATGGTAAATCCCAAGGCTAGGA-----CAGTGAACATGATTGACAAAGAAA--- 1692
; 523 LeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsn 542
; 1692 ----- 1692
; 543 AspValValLysPheTyrThrGluLeuIleGlyCysLysPheArgGlyIleArgIle 562
; 1693 -----ATGTTAATGAGCAAAA-----GTCACCTTCT 1719
; 563 GlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyr 582
; 1720 TGGACTTGGAAATTTAAGCCTCAACCTGCTATTTCGTTTCATCTCTTGTGCCCTGAACAT 1779

```

```
Qy 583 AlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaAlaThr 602
Db 1780 |||||
Qy 603 GluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle 622
Db 1798 |||||
Qy 623 IleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHisThr 642
Db 1837 |||||
Qy 643 IleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeu 662
Db 1897 |||||
Qy 663 ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAenAlaLys 682
Db 1945 |||||
Qy 683 LeuGlyGlyIleAenGlnLeuAspTrpSerGlyLeuIleAlaGluIleSerProGluGlu 702
Db 1990 |||||
Qy 703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
Db 2017 |||||
Qy 718 AspValThrHisProThrSerTyrSerGlyLeuAspTyrSerIleAlaAlaValValAla 737
Db 2077 |||||
Qy 738 SerIleAen---ProGlyGlyThrIleTyrArgAenMetIleValThrGln----- 753
Db 2137 |||||
Qy 754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
Db 2197 |||||
Qy 766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
Db 2257 |||||
Qy 786 GluAenAsnAspAenArgAlaProAlaHisIleValTyrArgAspGlyValSerAsp 805
Db 2299 |||||
Qy 806 SerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGln 825
Db 2350 |||||
Qy 826 PheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGln 845
Db 2410 |||||
Qy 846 LysArgHisAenThrArgLeuLeuArgArgMetGluLysAspLysProValAenLys 865
Db 2455 |||||
Qy 866 AspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAsp 885
Db 2500 |||||
Qy 886 MetLysGluSerLysGluThrGlyIleValAenProSerSerGlyThrThrValAspLys 905
Db 2509 |||||
Qy 906 LeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHisGlyValLeuGly 925
Db 2551 |||||
```

```
Qy 926 ThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGlu 945
Db 2611 |||||
Qy 946 ValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaAArgCysArgLysProIleSer 965
Db 2671 |||||
Qy 966 LeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArg 985
Db 2731 |||||
Qy 986 ThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThrArg 1001
Db 2779 |||||
RESULT 10
US-10-286-264-101
; Sequence 101, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddle, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Reuber, Lynne
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MEI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1149
US-10-286-264-101
Alignment Scores:
Pred. No.: 4e-42 Length: 2910
Score: 575.50 Matches: 237
Percent Similarity: 39.0% Conservative: 159
Best Local Similarity: 23.3% Mismatches: 337
Query Match: 10.8% Indels: 283
DB: 6 Gaps: 41
US-10-645-746-3 (1-1020) x US-10-286-264-101 (1-2910)
Qy 38 LysLysValLeuLeuLeuValAenTrpPheLysPheSerSerLysIleTyrAspArgGlu 57
Db 472 |||||
Qy 58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAenArgLysProGlyLysPro 77
Db 526 |||||
Qy 78 PheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
Db 570 |||||
Qy 98 ArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp 117
```

```
Db 570 ----- 570
Qy 118 ThrValTyrSerValCysArgLeuAenThrValThrSerLysMetLeuValSerGluLys 137
Db 571 -----ACAGTGAACAGAAACGTCATGAAACTTCTGGTTAAGAAT--- 609
Qy 138 ValValLysLysAspSerGlu-----LysLysAspGluLysAspLeuLysLys 154
Db 610 -----TATAAAGACTCTCATTGGGAGGAAGTACCAGCGTATGAT---GGAAGGAAA 660
Qy 155 IleLeuTyrThrMet-----IleLeuThrTyrArgLysLysPheHisLeuAsnPheSer 172
Db 661 AGCCTCTATACCTGCTGGTCCATTACCTTTTGACTCGAAAGAGTGTGTGTAATCGGCG 720
Qy 173 ArgGluAenProGluLysAspGluGluAlaAenArgSerTyrLysPhe---LeuLysAen 191
Db 721 GAGAAAAGAGCTGACGGTTCTCTCGGAAGGACAGACCGTTTAAAGTTGCTGTAAGAAT 780
Qy 192 Val-----MetThrGlnLysValArgTyrAla 200
Db 781 GTGACAAGCAGCTGATCTTTATACGTGTCACAGTTCCTTGATCGTAAGCAAGAGGCT 840
Qy 201 PropheValAsnGluGluLysValGlnPheAlaLysAsnPheValTyrAspAsnAen 220
Db 841 CCATAT-----GATACTATCCAGTCTGTGATGTTGTTCTTAGGGATAAGCCCTTAAT 894
Qy 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeu--- 239
Db 895 GATTATGCTCTGTTGGGAGGCTTTT-----TTCACACTAGTATTGGGA 939
Qy 240 -----GluValAlaProArgIleGluAlaTrpPheGlyIle 251
Db 940 AAGGACGCAAGAGATGTTAGGCGTGCCTGGAGATGGTATTGAGTACTGGAGAGGTTAT 999
Qy 252 TyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAsp 271
Db 1000 TTCAAAGTCTAAGG---CTGACTCAGATGGGTGTTGCTCTCAACACTGACGTTTCAGCA 1056
Qy 272 LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLysValAsp 291
Db 1057 AGATCATTTTAT-----GAA 1071
Qy 292 ProGlnSerCysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys 311
Db 1072 CCGATTGTTGTCACCTGATTTATTAGCAAGTTTCTGAATATAGGACTTA-----AAC 1125
Qy 312 MetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAenLeuLysLeu 331
Db 1126 AGGCCACTTAGAGACTCAGATCGACTTAAAGGTGAAGAAAGTTTTCAGGACACTGAAAGTT 1185
Qy 332 LysCysAlaGluValTrpAspAenGluMetSerArg----- 343
Db 1186 AAGTTGCTTCAC---TGGAAACGGCACAAAAGTGCACAAATTAGTGGGATTTCTAGTCTA 1242
Qy 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAenSerLeuValTyr 362
Db 1243 CCCATCAGGAGCTAAGTTTCACTCTGGAGGCAAAATCAGAAAGACGGTGTTCATAT 1302
Qy 363 LysValThrGlyLysSerAspArgGlyArgAenAlaLysLysLysTyrAspThrThrLeuPhe 382
Db 1303 -----TTTGCTGAAAATAATAAT----- 1320
Qy 383 LysIleTyrGluGluAenLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
Db 1321 -----GTGAAATACCAAGGCTCTACCTGCTATTCAAAACA 1359
Qy 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422
Db 1360 GGGAGTGACAAAGCCGCTCTACCTACCAATGGAGCTCTGCCAAATTCACGAA---GGG 1416
Qy 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
Db -----TCTGGTTTGATAAGGGAACATTTTCATAGCATTCAGGAGAGCT----- 2298
```

```
Db 1417 CAAGATACACCAAAAGGCTCAATGAGAAGCAAGTGACTGCTATTGCTAAAAGCTACCTGC 1476
Qy 443 ArgLysProHisAspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAspPheSer 462
Db 1477 CAACGACCCCTGAT---AGAGAGAACTCGATCAAAAACCTTGGTTGTGAAAAAATTAATAC 1533
Qy 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLysLeuGlnMetIleGlu 482
Db 1534 AATGATGATCTGAGC-----AAGGAGTTCGGATGTCAGTGACTACCCAACTAGCCTCG 1587
Qy 483 CysProGlyLysValLeuLysGluProMetLeuValAenSerValAenGluGlnIleLys 502
Db 1588 ATTGAAGCTCGTGACTTCCCCACCGATGTTGAAGTACCATGACAGTGGTAAAGAGAAA 1647
Qy 503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGlu 522
Db 1648 ATGTAAATCCAAAGGCTAGGA-----CAGTGAACATGATGTCACAGAAA--- 1692
Qy 523 LeuCysCysAlaValPheValAenGluThrAlaGlyAenProCysLeuGluGluAen 542
Db 1692 ----- 1692
Qy 543 AspValValLysPheTyrThrGluLeuIleGlyLysLysPheArgGlyIleArgIle 562
Db 1693 -----ATGTTTAAATGAGCAAAA---GTCACTTCT 1719
Qy 563 GlyAlaAsnGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAenGluTyr 582
Db 1720 TGGACTTCGGAATTTAAGCCTCAACCTGCTATTCCGTTTCATCTCTTGCCCTGGAACAT 1779
Qy 583 AlaPheTyrLysAenCysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaIleThr 602
Db 1780 -----ATTGAGGAAGCTCTTCTC 1797
Qy 603 GluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle 622
Db 1798 GAT-----ATCCAAAAAGGCGCCTGCTCTCCAA---CTGTTGATTT 1836
Qy 623 IleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAspHisThr 642
Db 1837 GTAATATTGCCGTGATGTGACTGGATCATATGGAATAATAAAGGATCTGTGAAACAGAA 1896
Qy 643 IleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeu 662
Db 1897 TTGGGGATTGCTCTCTCAGTGTGCCAACCTAGACAAGTTAATAAACTC----- 1944
Qy 663 ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAenAlaLys 682
Db 1945 -----AACAAAGCAGTACATGGAATAATGTTGCCCTTGAAGATCAATGTCAAG 1989
Qy 683 LeuGlyGlyIleAenGlnGluLeuAspTrpSerGluIleAlaGluLysSerProGluGlu 702
Db 1990 ACTGGGGAAGGAACACTGTTCTTAAT----- 2016
Qy 703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
Db 2017 GATGCTATTAGAAGAAACATACCTCTTATTACTGATCGTCCAAACCATCATCATGGGTGT 2076
Qy 718 AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaIleValAla 737
Db 2077 GATGTGACTACCCACAGCCTGGAGAGGACTCAAGTCCCTTCTTATGCTGCTGTGTGGCC 2136
Qy 738 SerIleAen---ProGlyGlyThrIleTyrArgAenMetIleValThrGln----- 753
Db 2137 TCTATGGAGCTGGCTTGAGATAAACAAATACCGAGATTTGGTTCTGCTCAAGCTCATAGG 2196
Qy 754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
Db 2197 GAAGAAATTTATTCAGGACCTGTATAAGCTGGTTTCAGGATCCACAAAGCTGGGTAGTCCAC 2256
Qy 766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
Db 2257 -----TCTGGTTTGATAAGGGAACATTTTCATAGCATTCAGGAGAGCT----- 2298
```


QY	221	SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeu---	239	QY	563	GlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyr	582
DB	895	GATTATGTCCTCTGTTGGGAGGCTTTT---	939	DB	1720	TGGACTTGGGAATTTAAGCCTCAACTGCTATTTCGTTTCATCTCTGTGTCCTCCCTGAACAT	1779
QY	240	-----GluValAlaProArgIleGluAlaTrpPheGlyIle	251	QY	583	AlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluGluLeuAlaThr	602
DB	940	AAGGACGCAAGAGATCGTAGGGGTGAGCTTGGAGATGGTATTGAGTACTGGAGAGGTTAT	999	DB	1780	-----ATTGAGGAGAGCTCTCTC	1797
QY	252	TyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAsp	271	QY	603	GluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle	622
DB	1000	TTCCAAAGCTTAAGG---CTGACTCAGATGGGTTTGTCTCTCAACATTCAGCTTTTCAGCA	1056	DB	1798	GAT-----ATCCACAAAGGCGCCTGCTCTCCAA-----CTGTGTGATT	1836
QY	272	LysLeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAsp	291	QY	623	IleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAspHisThr	642
DB	1057	AGATCATTTTAT-----GAA	1071	DB	1837	GTAATATTGCCGTGATGCTGATCATATATGAAAAATAAAGGATCTGTGNAACAGAA	1896
QY	292	ProGlnSerCysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLys	311	QY	643	IleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeu	662
DB	1072	CCGATTGTTGTCACCTGACTTTATTAGCAAGTTTCTGAATATAAGGGACTTA-----AAC	1125	DB	1897	TTGGGGATTGCTCTCTCAGTTTGCACCACTAGACAAGTTAATAAACTC-----	1944
QY	312	MetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeu	331	QY	663	ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLys	682
DB	1126	AGGCCACTTAGAGACTCAGATCGACTTAAGGTGAAGAAAGTTTTCAGGACACTGAAAGTT	1185	DB	1945	-----AACAGCAGTACATGGAATAATGTCCTTGAAGATCAATGTCAG	1989
QY	332	LysCysAlaGluValTrpAspAsnGluMetSerArg-----	343	QY	683	LeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGluIleSerProGluGlu	702
DB	1186	AGTTCCTTCAC---TGGAACGGCACAAAAGTGCCTCAAAATAGTGGGATTTCTAGTCTA	1242	DB	1990	ACTGGGGNAGGAACACTCTTCTTAAT-----	2016
QY	344	---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyr	362	QY	703	LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle	717
DB	1243	CCCATCAGGAGCTAAGGTTCACTCTGGAGGACAAATCAGAGAAGCGTGTGTCATAT	1302	DB	2017	GATGCTATTAGAAGAAACATACCTCTTATTACTGATCGTCCAAACCATCATCTGGGTGCT	2076
QY	363	LysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPhe	382	QY	718	AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAla	737
DB	1303	-----TTTGCTGAAAAATAATAAT-----	1320	DB	2077	GATGATGCTACCCACAGCCTGGAGAGACTCAAGTCTTCTATTGCTGCTGTGTGGCC	2136
QY	383	LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal	402	QY	738	SerIleAsn---ProGlyGlyThrIleTyrArgAsnMetIleValThrGln-----	753
DB	1321	-----TATAGA-----GTGAATACCAAGCTCTACCTGCTATTCAACA	1359	DB	2137	TCATGAGTGCCTCGCTGAGATAACAATAACCGAGGATTGGTTCTGCTCAAGCTCATAGG	2196
QY	403	LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro	422	QY	754	GluGlu-----CysArgProGlyGluArgAlaValAlaHis	765
DB	1360	GGGAGTGACACAAGACCCGCTTACCTACCCTAATGGAGCTCTGCCAAATTCACGAA---GGG	1416	DB	2197	GAAGAAATTTATTCAGGACCTGTATAGCTGGTTCCAGGATCCACACGCTGGGTAGTCCAC	2256
QY	423	GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr	442	QY	766	GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla	785
DB	1417	CAAGATACACCAAAAGGCTCAATGAGAAAGCAAGTGACTGCATTTGCTAAAGCTACCTGC	1476	DB	2257	-----TCTGGTTTGATAGGGACATTTTCATAGCATTCAGGAGAGCT-----	2298
QY	443	ArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSer	462	QY	786	GluAsnAsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAsp	805
DB	1477	CAACGACCCCTGAT---AGAGAGAACTCGATCAAAAACCTTGGTTGTGAAAAATAATTAC	1533	DB	2299	-----ACAGGCCAGATACCTCAAGGATCATCTTCTATCGTCGACGAGTAAGCGAA	2349
QY	463	SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu	482	QY	806	SerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGln	825
DB	1534	AATGATGATCTGAGC-----AAGGAGTTTGGGATGTCAGTGACTACCCCACTAGCCTCG	1587	DB	2350	GGGCAGTATTAGTCAGGTTCTGCTACATGAGATGACTGCTATCCGCAAGCCTTGTAACCT	2409
QY	483	CysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLys	502	QY	826	PheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGln	845
DB	1588	ATTGAAGCTCGTGTACTTCCCCACCGATGTTGAAGTACCATGACAGAGTGTTAAAGAGAA	1647	DB	2410	CTCCAA-----GAGAAATATGTTCTCTCGTGTGTACTTTCGTGTGTGTCAG	2454
QY	503	MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValProGluLysGlu	522	QY	846	LysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLys	865
DB	1648	ATGGTAATCCAAAGGCTAGA-----CAGTGGAACTAGTTGACACAGAAA---	1692	DB	2455	AAACGTCACACACACGTTTGTTCCTGACCAACACGGG-----AATCGT	2499
QY	523	LeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluAsn	542	QY	866	AspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAsp	885
DB	1692	-----	1692	DB	2500	GATATGACT-----	2508
QY	543	AspValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheArgGlyIleArgIle	562	QY	886	MetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLys	905
DB	1693	-----ATGGTTAATGGAGCAAA-----GTCATCTCT	1719	DB	2509	-----GATAAGAGTGGCAATATTCAACCA-----GCTACTGCTGGGACACT	2550
				QY	906	LeuIleValSerLysTyrLysPheAspPheLeuAlaSerHisGlyValLeuGly	925

Db 1072 CCGATTGTTGTCACGTCTTATTAGCAAGTTTCTGAATATAAGGGACTTA-----AAC 1125
QY 312 MetThrIleArgGlnAlaalaArgProArgIleArgGlnLeuLeuLeuLeuLeuLeuLeu 331
Db 1126 AGGCCACTTAGAGACTCAGACTCAGACTTAAGGTGAAGAAAGTTTTCAGGACACTGAAGTT 1185
QY 332 LysCysAlaGluValTrpAspAsnGluMetSerArg----- 343
Db 1186 AAGTGTCTTCAC---TGGAAACGGGCACAAAAGTCCCAAAATAGTGGGATTTCTAGTCTA 1242
QY 344 ---LeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluGluAsnSerLeuValTyr 362
Db 1243 CCCATCAGGAGCTAAGGTTCACTCTGGAGGACAAATCAGAGAACGGTGTGTCAATAT 1302
QY 363 LysValThrGlyLysSerAspArgGlyArgAsnAlaLysGlyTyrAspThrThrLeuPhe 382
Db 1303 -----TTTGCTGAATAATAAT----- 1320
QY 383 LysIleTyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysVal 402
Db 1321 -----TATAGA-----GTGAATACCCAGGCTCTACCTCTCTATTCAACA 1359
QY 403 LysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysPro 422
Db 1360 GGGAGTGACAAAGACCCCTCTACCTACCAATGGAGCTCTGCCAAATTGACGAA---GG 1416
QY 423 GlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr 442
Db 1417 CAAAGATACACCAAAAGGCTCAATGAGAAAGCAAGTGCATTTGTAAAGCTACCTGC 1476
QY 443 ArgLysProHisAspTyrLysGluLeuThrLeuLysMetLeuLysGluLeuAspPheSer 462
Db 1477 CAACGACCCCTGAT---AGAGAGAACTCGATCAAAAACCTTGTTGTGAAAAATTAATAC 1533
QY 463 SerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGlu 482
Db 1534 AATGATGATCTGAC-----AAGGAGTTTGGGATGTGAGTGCAGTACCCCACTAGCCTCG 1587
QY 483 CysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLys 502
Db 1588 ATTGAAGCTCGTGTACTTCCCCACCGGATGTTGAAGTACCACGACAGTGTGTAAGAGAAA 1647
QY 503 MetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGlu 522
Db 1648 ATGGTAAATCCAAAGCTAGGA-----CAGTGAACATGATTGACAAAGAA--- 1692
QY 523 LeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsn 542
Db 1692 ----- 1692
QY 543 AspValValLysPheTyrThrGluLeuIleGlyGlyCysLysPheArgGlyIleArgIle 562
Db 1693 -----ATGGTTAATGGAGCAAAA-----GTCACTTCT 1719
QY 563 GlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyr 582
Db 1720 TGGACTTGGAAATTTAAGCTCAACCTGCTATTCCTGTTTCCCTCCCTGACAT 1779
QY 583 AlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaAlaThr 602
Db 1780 -----ATTGAGGAGCTCTTCTC 1797
QY 603 GluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle 622
Db 1798 GAT-----ATCCAAAAAGGGCACCTGGTCTCCAA-----CTGTTGATT 1836
QY 623 IleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAspHisThr 642
Db 1837 GAAATATCCCTGATGTGACTGATCATATGGAAAAATTAAGGATCTGTGAAACAGAA 1896
QY 643 IleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAsnSerLeu 662
Db 1897 TTGGGGATTGTCTCTCAGTGTGTCACCACTAGACAAAGTTTAATAAACTC----- 1944

QY 663 ArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLys 682
Db 1945 -----AACAGCAGTACATGGAAATATGTTGCTTGAAGATCAATGTCAAG 1989
QY 683 LeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGluIleSerProGluGlu 702
Db 1990 ACTGGGGAGNAACACTGTTCTTAAT----- 2016
QY 703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
Db 2017 CATGTCTATTAGAAAGAACATACCTCTATTACTGATCGTCCCAACCATCATGGGTGCT 2076
QY 718 AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValAlaAla 737
Db 2077 GATGTGATCCACCCACGCTCGAGAGACTCAAGTCTCTTATTGCTGCTGTGTGGCC 2136
QY 738 SerIleAsn---ProGlyGlyThrIleTyrArgAsnMetIleValThrGln----- 753
Db 2137 TCTATGGACTGCGCTGAGATAACAATACCGAGGATTGGTTTCTGCTCAAGCTCATAG 2196
QY 754 GluGlu-----CysArgProGlyGluArgAlaValAlaHis 765
Db 2197 GAAGAAATTAATTCAGGACCTGTATAAGCTGGTTTCAGGATCCCAACGCTGGGTAGTCCAC 2256
QY 766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
Db 2257 -----TCTGGTTTGATAAGGAAACATTTTCATAGCATTCAGGAGACT----- 2298
QY 786 GluAsnAsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAsp 805
Db 2299 -----ACAGGCCAGATACCTCAAGGATCATCTTATCGTGACGAGTAAGCGAA 2349
QY 806 SerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGln 825
Db 2350 GGGCAGTTTAGTCAGGTTCTGCTACATGAGATGACTGCTATCCGCAAGCTTGTAACTCT 2409
QY 826 PheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGln 845
Db 2410 CTCCAA-----GAGAAATATGTTCTCTCGTGTTACTTTCGTGATTGTCAG 2454
QY 846 LysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLys 865
Db 2455 AAACGTCACACACACGTTTGTCTCTGAGCAACACGGG-----AATCGT 2499
QY 866 AspLeuThrProAlaGluThrAspValAlaValAlaLysGlnTrpGluGluAsp 885
Db 2500 GATATGACT----- 2508
QY 886 MetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLys 905
Db 2509 -----GATAAGTGGCAATATTCAACCA-----GGTACTGCTGGACACT 2550
QY 906 LeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHisHisGlyValLeuGly 925
Db 2551 AAAATCTGTCCACCTAATGAATTCGACTTCTATTTGAACAGCCATCTCTGTATTCAGGA 2610
QY 926 ThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGlu 945
Db 2611 ACAAGCGCCGCGCACATTACCATGTACTTCTCGATGAGAACGGTTTCCCGCTGATCAG 2670
QY 946 ValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSer 965
Db 2671 TTGCAATGTCTCAACACACCTCTGTACAGTATGCGAGGTGTACAAAATCTGTGTCA 2730
QY 966 LeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArg 985
Db 2731 ATTGTGCCACCGCCTACTACGCTCCTTGGCTGCATTCGGTGC-----CGC 2778
QY 986 ThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThrArg 1001
Db 2779 TACTACATGGAGAGTGCATGCTGTGAGGAGTTCGAGCAGCGGTCCAGG 2826

Db 1443 CAGGGCGTCTGGGACATCGGACAGCAGTTCACACGGGCGATCGAGATCAAGGTGTGG 1502
Qy 524 CybCybAlaValPheValValAenGluThrAlaGlyAsnProCybLeuGluAenAsp 543
Db 1503 GCCATTGCGTCTCGCCCCCAG-----CGCCAGTCAGCAGGAAGTCCAT--- 1547
Qy 544 ValValLysPheThrThrGluLeuLeu-----GlyGly-CysLysPhe 557
Db 1548 ---CTGAAGTCTTTCACAGCAGCTCAGAAAGATCTCGAGAGACGCCGCGCATGCCATC 1604
Qy 557 eArgGlyLeuArgGlyLeuAenGluAenArgGlyAlaGlnSerIle-----MetTy 575
Db 1605 CAGGGCCAGCGTGTCTTCTGCAATACGCG-CAGGGGGCGGACAGCGTGGAGCCCATGTT 1663
Qy 575 rAspAlaThrLysAenGluThrAlaPheTyLysAenCysThrLeuAenThrGlyIleG1 595
Db 1664 CCGGCACCTGAAGACACGTATGCG----- 1688
Qy 595 yArgPheGluIleAlaAlaThrGluAlaLysAenMetPheGluArgLeuProAspLysG1 615
Db 1688 ----- 1688
Qy 615 uGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAenAlaTyxGlyPheVa 635
Db 1689 ---GGCTGCGAGCTGTGTGTGTATCTCTCGCGCAAGACGCCGCGTGTACGCCAGGT 1744
Qy 635 llybHisTyxAspHisThrIleGlyValAlaAenGlnHisIleThrSerGluThrVa 655
Db 1745 CAAGCGCGGGAGACACGGTGTGGGATGCCACGACGTCGTCGAGATGAAGACGT 1804
Qy 655 lThrLysAlaLeuAseLeuArgHisGluLysGlySerLysArgIlePheTyxGlnI1 675
Db 1805 G-----CAGAGGACACGCGCACAGACCCCTGTCCAACCT 1837
Qy 675 eAlaLeuLysIleAenAlaLysLeuGlyGlyIleAenGlnGluLeuAspTrpSerGluI1 695
Db 1838 CTGCTGAAGATCAAGCTCAAGCTGGAGCGGTGAACAACATCTCTG----- 1883
Qy 695 eAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyxVa 715
Db 1884 -----CTGCCCGGCGAGCGCGCGGTGTTCACGACGCGCGTC---ATCTTTCT 1930
Qy 715 lGlyLeuAspValThrHisProThrSerTyxSerGlyIleAspTyxSerIleAlaAlaVa 735
Db 1931 GGGAGCAGACGTCACCTACCCCGCGCGGATGGGAAGAGCCCTCCATTGCGCGCGT 1990
Qy 735 lValAlaSerIleAenProGlyGlyThrIleTyxArgAenMetIleValThrGlnGluG1 755
Db 1991 GTTGGCAGCAGTGGAGCCCGCCCGATCTGCGCCCGCGTGTGTGAGCAGC--- 2048
Qy 755 uCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
Db 2049 -----CACCAGGAGAGATCATACAGACCTGCGCCG 2080
Qy 775 aLysPheValLysLeuLeuArgGluPheAlaGluAenAenAspAenArgAlaProAlaHi 795
Db 2081 CATGTCTCCGAGCTCTCATCCAGTTCTACAGTCCACGCGCTTCAAG---CCACCCCG 2137
Qy 795 sIleValValTyxArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG1 815
Db 2138 CATCATCTTCTACCGGACGGGTCTCTGAAGCCCGATTCACAGAGTTCTCCACACGA 2197
Qy 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
Db 2198 GTTGTGGCCATCCGTGAGCGCTGTATCAAGTA-----GAAAGAGACTA 2242
Qy 835 oGluProLysTyxThrPheIleValIleGlnLysArgHisAenThrArgLeuLeuArgAr 855
Db 2243 CCAGCCCGGGATCACCTTCTATCGTGTGAGAGAGGACACACCCCGGCTCTTC---TG 2299
Qy 855 gMetGluLysAspLysProValAlaAenLysAspLeuThrProAlaGluThrAspValAl 875

Db 2300 CACTCACAAAGACGAGCGGGTGGGAAAAGTGGAAACATTCACGCA----- 2345
Qy 875 aValAlaAlaValLysGlnTrpGluLeuAspMetLysGluSerLysGluThrGlyIleVa 895
Db 2345 ----- 2345
Qy 895 lAenProSerSerGlyThrThrValAspLysLeuIleValSerLysTyxLysPheAspPh 915
Db 2346 ---GGCAGACTGTGGACACGAAATCACCCACCCACCGAGTTCGACTT 2392
Qy 915 ePheLeuAseHisHisGlyValLeuGlyThrSerArgProGlyHisTyxThrValMe 935
Db 2393 CTACTGTGTAGTCAGCTGGCATCCAGGGGACAGAGCGCTTCGCACTATCACGTCCT 2452
Qy 935 tTyxAspLysGlyMetSerGlnAspGluValTyxLysMetThrTyxGlyLeuAlaPh 955
Db 2453 CTGGACGACAAATCTCTCTCTGTGATGAGTGTGAGATCTTAACCTACCGCTGTGTCA 2512
Qy 955 eLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyxAlaHisLe 975
Db 2513 CACTACGTGGCTGCACACGCTCGTGTCCATCCAGCGCCAGCATACTAGCTCACCT 2572
Qy 975 uSerCysGluLysAlaLysGluLeuTyxArgThrTyxLysGluHis 990
Db 2573 GGTGGCTTCGGGGCCAGG---TACCACCTGTGTGATAAGGAACAT 2615
RESULT 14
US-10-467-397-17
; Sequence 17, Application US/10467397
; Publication No. US20040137448A1
; GENERAL INFORMATION:
; APPLICANT: THORNTON, Michael; HAPALIA, April J. A.;
; APPLICANT: LU, Dying Aina M.; ARVIZU, Chandra S.;
; APPLICANT: SWARNAKAR, Anita; LU, Yan;
; APPLICANT: WARREN, Bridget A.; BAUGHN, Mariah R.;
; APPLICANT: TANG, Y. Tom; LEE, Ernestine A.;
; APPLICANT: YAO, Monique G.; RAMKUMAR, Javalaxmi;
; APPLICANT: KHAN, Farrah A.; GANDHI, Ameena R.;
; APPLICANT: DING, Li; YUE, Henry;
; APPLICANT: GIETZEN, Kimberly J.; CHAWLA, Narinder K.;
; APPLICANT: THANGAVELU, Kavitha; ELLIOTT, Vicki S.
; APPLICANT: MARQUIS, Joseph P.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PI-0368 USN
; CURRENT APPLICATION NUMBER: US/10/467,397
; PRIOR FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US02/03844
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/268,118
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/270,963
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,858
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/271,194
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,071
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/283,496
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/344,650
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1725129CB1
; FEATURE:
; NAME/KEY: unsure

LOCATION: 3528
OTHER INFORMATION: a, t, c, g, or other
US-10-467-397-17

Alignment Scores:
Pred. No.: 4,1e-41 Length: 3580
Score: 566.00 Matches: 239
Percent Similarity: 40.0% Conservative: 159
Best Local Similarity: 24.0% Mismatches: 385
Query Match: 10.6% Indels: 214
DB: 8 Gaps: 38

US-10-645-746-3 (1-1020) x US-10-467-397-17 (1-3580)

```

Qy 19 ProGluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
Db 95 CCCATCCAAAGATATGCTTCAAGCTCCACCTAGACCCGACTTGGGACCTCCGGGAGA 154

Qy 39 LysValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyr 58
Db 155 ACAATCAATACAGGCCAATTCCTCGAATGGACATCCCAAAATGGACATCATCAT 214

Qy 59 TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhe 78
Db 215 TATGAATGGATATCAAGCCAGAGAG-----TGC 244

Qy 79 ProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArg 98
Db 245 CCGAGGAGATTAAACAGGGAATCGTGGACACATG-----GTCCAG 286

Qy 99 HisGluLysLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLys 116
Db 287 CACTTTAAACACAG-----ATCTTTGGGATCGGAAGCCGCTTGAACGGCAGG 337

Qy 117 AspThrValTyrSerValCysArgLysLeuAsnThrValThrSerLysMetLeuValSerGlu 136
Db 338 AAGAAATCTATACAGCCATCCCTTCCGATGGGAGGCAAGGTG-----GAG 398

Qy 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeu 156
Db 389 CTGGAGTCAAGCTGCCAGGAGAGGCAAGGATCG-----ATCTTCAAGGTGTCCATC 442

Qy 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
Db 443 AAGTGGGTGCTCGTGAGCTTGAGCGCTTACAGATGCATTTTCAGGCGCGCTGCC 502

Qy 177 GluLysAspGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln--- 195
Db 503 AGCGTCCCTTTTCAG-----ACGATCCAGGCCCTGGACGTGTGATGAGGCATCTG 553

Qy 196 ---LysValArgTyrAlaProPheValAsnGluLysValGlnPheAlaLysAsn 214
Db 554 CCATCCATGAGGTACACCCCGTG----- 577

Qy 215 PheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg 234
Db 578 -----GGCCGCTCTTTCACCGCTCCGAAGGCTGCTTAACCT----- 619

Qy 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGly 254
Db 620 -----CTTGGCGGGGCGGAGAGTGTGGTTTGGCTTCCATCATGCTCC 661

Qy 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
Db 662 GTCCGCTCTCTCTC---TGGAAATGATGCTGAATATTGATGTGTGTCAGCAACAGCGTTT 718

Qy 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSer 294
Db 719 TACAAGGCA-----CAGCCAGTAATCGAGTTTGTGTAAGTTTGGATTTTAAAGT 772

Qy 295 CysAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
Db 773 ATTGAAGACAA-----CAAAACCTCTG 796

```

```

Qy 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys----- 332
Db 797 ACAGATTCCCAAGGGTAAAGTTTACCAAGAAATTAAGGCTTAAGAGGTGAGATAACG 856

Qy 333 ---CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGluArg 347
Db 857 CACTGTGGGAGATGAAGAGAAAGTACCGGCTCTGCAATGTGACCCCGCGCCCGCCAGT 916

Qy 348 HisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLys 367
Db 917 CACCAACATTC---CCGCTGCAGCAGAGCGGCGGAGCGGTGGAGTGCAGC----- 967

Qy 368 SerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGlu 387
Db 968 -----GTGGCCAGTAT-----TTCAAG-----GACAGG 991

Qy 388 AsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLys 407
Db 992 CACAAGTTGGTCTTGGCGCTACCCCACTCCCATGTTTACAAGTCGACAGGACAGAAA 1051

Qy 408 GluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsn 427
Db 1052 CACACCTACCTTCCCTGGAGGTCTGTAAACATTTGTGCAGCACAAGATGTATTAAAAA 1111

Qy 428 ArgIleAspLeuValMetGlnAspLysPheLysLysArgAlaThr---ArgLysProHis 446
Db 1112 TTAACGGAC-----AATCAGACCTCAACCATGATCAGACGAGCTGCTAGTGGCGGCC 1165

Qy 447 AspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeu 466
Db 1166 GATCGGCAAGAGAGATTAGCAAAATTTGATCGAAGTCAAGTTTCAACACAGATCCCA 1222

Qy 467 AsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLys 486
Db 1223 ---TAGTCCGTGAATTTGGAATCATGCTCAAGATGAGATGACAGACGTGACTGGCGG 1279

Qy 487 ValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal 506
Db 1280 GTGCTGCAGCGCGCTCCATCTCTACGGGGCGAGGAATAAAGCTATTATGCGACCCCTGTC 1339

Qy 507 -----IleArgGlyPheGlnGluLysGlnLeuAsnVal---ValProGluLysGluLeu 523
Db 1340 CAGGGCGCTGCGACATCGCGAAACAAGCATCTCCACAGCGGCATCGAGATCAAGGTGTGG 1399

Qy 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysValLeuGluAsnAsp 543
Db 1400 GCCATTGCGTGTCTCGCCCCCAG-----CGCCAGTGCACGGAGTCCAT--- 1444

Qy 544 ValValLysPheTyrThrGluLeuIle-----GlyGly-CysLysPh 557
Db 1445 ---CTGAAGTCTTTCACAGAGCAGCTCAGAAAGATCTCGAGAGACGCTGCGCATGCCCATC 1501

Qy 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIle-----MetTyr 575
Db 1502 CAGGGCCAGCCGCTGCTTCTGCAAAATACGCG-CAGGGGGCGGACAGCGGTGGAGCCCATGTT 1560

Qy 575 xAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleG 595
Db 1561 CCGGCACCTCGAAGAACACGTATGCG----- 1585

Qy 595 yArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysG 615
Db 1585 ----- 1585

Qy 615 uGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheVa 635
Db 1586 ----GSCCTGCAGCTGGTGGTGTCTCTCGCCCGCAAGACGCCCGGTGTACCCGAGGT 1641

Qy 635 lLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVa 655
Db 1642 CAAAGCGGTGGGAGACACCGTGTGGGGATGGCCACGACGTGCTGCAGATGAAGACGT 1701

```

QY 655 lThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIrl 675
 Db 1702 G-----CAGAGGACCAGCCAGCCAGACCCCTGTCCAACT 1734
 QY 675 eAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTyrSerGluIrl 695
 Db 1735 CTGCTGAAGATCAACGTCAAGCTGGAGCGGTGAACAACATCTTG----- 1780
 QY 695 eAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrVa 715
 Db 1781 -----CTGCCCCAGGCGAGCGCGCGGTGTTCAGCAGCCGCTC---ATCTTTCT 1827
 QY 715 lGlyIleAspValThrHisGlyThrSerTyrSerGlyIleAspTyrSerIleAlaIle 735
 Db 1828 GGAGCAGACGTCATCAACCCCGCGCGGATGGGAAGAAGCCCTCCATTTGCCGCGCT 1887
 QY 735 lValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluGl 755
 Db 1888 GGTGGCAGCATGGACGCCCAACCAATCGCTACTGCGCCAGCGTGGCGTGCAGCAG-- 1945
 QY 755 uCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
 Db 1946 -----CACCGCAGGAGATCATACAAGACCTGGCCGC 1977
 QY 775 aLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAlaHi 795
 Db 1978 CATGTCTCCGAGCTCTCATCGATTCTACAAGTCCAGCGCTTCAAG---CCACCCG 2034
 QY 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGl 815
 Db 2035 CATCATCTTCTACCGCGAGCGGTCTCTGAGGCCAGTTCACGAGTTCTCCACACCA 2094
 QY 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
 Db 2095 GTTGTCTGGCATCCGTGAGCGCTGTATCAAGCTA-----GAAAAAGACTA 2139
 QY 835 oGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuArgAr 855
 Db 2140 CAGCCCGGGATCACCTTCTATGTGTGAGAGAGGACACACCCCGCTCTC---TG 2196
 QY 855 gMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAl 875
 Db 2197 CACTGACAAAGACGCGGTGGGAAAGTGAACATTCAGCA----- 2242
 QY 875 aValAlaAlaValLysGlnTrpGluAspMetLysGluSerLysGluThrGlyIleVa 895
 Db 2242 ----- 2242
 QY 895 lAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPh 915
 Db 2243 -----GGCACACCTGTGGACACGAAATCACCCACCCACCGAGTTCGACTT 2289
 QY 915 ePheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
 Db 2290 CTACCTGTGTAGTCACGCTGGCATCCAGCGGACAGCGCTTCGCACTATCACGTCCT 2349
 QY 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955
 Db 2350 CTGGGACGCAATCGTTTCTCTCTGATGCTGCAGATCTTAACCTACACGCTGTGTC 2409
 QY 955 eLeuSerAlaAtyArgLysProIleSerLeuProValProValHisTyrAlaHisLe 975
 Db 2410 CACCTAGTGGCTGCACAGCTCGTGTTCATCCAGCGCCAGCATACATACGCTCACCT 2469
 QY 975 uSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
 Db 2470 GGTGGCTTCCGGGCCAGG---TACCACCTGTGGGATAAGGAACAT 2512

RESULT 15

US-11-097-143-27926
 ; Sequence 27926, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; TITLE OF INVENTION: DROSOPHILA GENES.
 ; FILE REFERENCE: CL000728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; CURRENT FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 27926
 ; LENGTH: 3486
 ; TYPE: DNA
 ; ORGANISM: DROSOPHILA
 ; US-11-097-143-27926

Alignment Scores:

Pred. No.: 5,446-41 Length: 3486
 Score: 564.50 Matches: 233
 Percent Similarity: 40.4% Conservative: 174
 Best Local Similarity: 23.1% Mismatches: 394
 Query Match: 10.6% Indels: 206
 DB: 13 Gaps: 37

US-10-645-746-3 (1-1020) x US-11-097-143-27926 (1-3486)

QY 38 LysLysValLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGlu 57
 Db 527 CGCCCGATTGTCTGCGGCCAATCCTTCACAGTGACA-----ATGCGCGTGGC 577
 QY 58 Tyr---TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLys 76
 Db 578 TATGTGCATCACTATGACATCAATATACAGCCGACAGTGTCCGCGAAAGGTGAACCGT 637
 QY 77 ProPheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHis 96
 Db 638 GAGATT-----ATCGAGACTATGTGTGCATGCTTATAGCAAGATCTTCGA----- 682
 QY 97 LeuArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLys 116
 Db 683 -----GTGCTCAACCGC---GTGTTTCGATGTCGCG 709
 QY 117 AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
 Db 710 AACATCTGTATACCGCGGATCCCTGCCCATTTGGCAACAGCGGTCTA-----GAG 760
 QY 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeu 156
 Db 761 CTGGAGGTTACTCTTACCCGCGCGGCGGCAAGATCGA----- 796
 QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
 Db 797 -----ATCTTTCCGCTGACGATCAAGTGGCAGGCTCAGGCTCTCGCTCTCAATTG 847
 QY 177 GluLysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLys 196
 Db 848 GAGGAAGTCTCTCGAAGCGCGC-----ACGCGCGAC 877

QY	197	ValArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheVal	216
DB	878	ATACCTATGATGCATTTTGGCGTTCGATGTGGTCATCGCGCATCTCCCGCAGCATGACG	937
QY	217	TyrAsp-----AsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg	234
DB	938	TACACGCCAGTGGGAGCTAGCTTCTTCAGTTCCTCCCGAGGGTTACTACATCCC-----	991
QY	235	PheGluGlnSerLeuGluValAlaProArgIleGluAlaTyrPheGlyIleTyrIleGly	254
DB	992	-----CTGGTGGTGGACGCGAGGTTTGGTTTCGTTTCCATTCATCC-----	1033
QY	255	IleLysGluLeuPheAspGlyGluProValIleAsnPheAlaIleValAspLysLeuPhe	274
DB	1034	GTAGGCC---TCCAGTGGAGAGATGCTCAATATCGAGTCTCGGCCACCGCTTTC	1090
QY	275	TyrAsnAlaProLysMetSerLeuLeuAspTyrIleLeuLeuLeuIleValAspProGlnSer	294
DB	1091	TACAAGGCT-----CAACCAAGTCACTTCATGTCCGAGGTGTCGACATTCGCGAC	1144
QY	295	CysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle	314
DB	1145	ATCAACGACGAG---CGCAAAACCGCTC-----	1168
QY	315	ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLysCysAla	334
DB	1169	ACCATTCCGACGCGCTCAAGTTCACGAAGAGATCAAGGCTTTGAAGATCGAGATCACC	1228
QY	335	GluValTyrAspAsnGlnMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeu	354
DB	1229	CAC-----	1231
QY	355	CysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArgGly-----	371
DB	1232	TCCGCCAGATGCGTCGCAAGTATCGTGTGTGCAACGTCATCTGCCGCCCGCTCAGATG	1291
QY	372	-----ArgAsnAlaLysLysTyrAspThrThrLeuPheLysIle	384
DB	1292	CAATCATCCCACTGCAGCTGGAGAACGGACAGCCGTAGAGTGCACCGTGGCCAGATAC	1351
QY	385	Tyr---GluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLys	403
DB	1352	TTCCTGGACAGTACCGCATGAAATTGGCTACCGCACTTGCCCTGCCTGCAGTTGGC	1411
QY	404	SerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGln	423
DB	1412	CAAGAGCACACACACTTACTTCCTCTAGAGGTGTGCAACATT---GTGCCCGGACAG	1468
QY	424	ArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArg	443
DB	1469	CGGTGCATTAAAGCTGACCGATATGCAGAGCTGCACCATGATCAAGGCCACAGCTCGT	1528
QY	444	LysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSer	463
DB	1529	TCAGCTCCGGATCGTAGAGCGTGAGATTAACAACCTTGGTAAAGCGCGCGACTTCAACAC	1588
QY	464	GluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCys	483
DB	1589	GAT-----TCGTATGTGCAAGATTTGGCTCGACCATCTCCATTTCGATGTGAGGTA	1642
QY	484	ProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMet	503
DB	1643	CGAGAGCGGCTCTTGCTCTCCCAAGCTTACGTATGGGGACGT-----	1687
QY	504	ThrProValIleArgGlyPheGlnGlnLysGlnLeuAsnValValProGluLysGluLeu	523
DB	1688	-----GTGTCTACCGGCTCACCAGCGCAGCAGCTG---TTCGCCGCCACAGAACAGGTG	1738
QY	524	CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp	543
DB	1739	AGCTTGCCC-----TCGCCCAACAGGGGTGTATGGGATATGCGA	1777

Qy	544	ValVallylsPheTyrThrGluLeuLeuIleGlyCysLysPheArgGlyIleIleArgGlyGly	563
Db	1778	GCCAAGCAGTGTCTTCACTGGCGTC-----GAGATCCGCATCTCG	1816
Qy	564	AlaAsnGluAsnArgGlyAlaGln---SerIleMetTyrAspAlaThrLysAsnGluTyr	582
Db	1817	GCCATCGCCTGTTTGGCCCAACGCGGATCGCGAGGATCGCTGCGTAATATTCACC	1876
Qy	583	AlaPheTyrLysAsnCysThrLeuAsnThrGly-----IleGly-----	595
Db	1877	CAGCAGCTGCAGAGATCTCAACAGATGCAGGCATGCCGATAATTGGACAGCGTGCCTC	1936
Qy	596	---ArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLys	614
Db	1937	TGTAAGTACGCCACCGCGCGCGATCAAGTGGAAACCCATGTTCCGTTACCTGGAAGATCACC	1996
Qy	615	GluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPhe	634
Db	1997	TTCCCCGGCCTGCMACTCTGTCGTGTGTCTGCCCGCAAGACTCCAGTATACGCCGAG	2056
Qy	635	ValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThr	654
Db	2057	GTGAAGCGTGTAGTGTACACCGTCTCGGTATGCGCACCCAGTGTGTGCAGGCCAAGAAC	2116
Qy	655	ValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGln	674
Db	2117	GTGAACAAGACA-----TCGCCACAGACGCTCTCTAAT	2149
Qy	675	IleAlaLeuLysIleAsnAlaLysLeuLysGlyIleAsnGlnLutLeuAspTrpSerGlu	694
Db	2150	CTGTGCTGAAGATCAACGTCAAAGTGTGGCGGATCAATTCATTCGTG-----	2197
Qy	695	IleAlaGluIleSerProGluGluLysGluArgLysThrMetProLeuThrMetTyr	714
Db	2198	GTCCCTCCATTCGGCCAAAGTCTTCATGACCG-----GTCATCTTT	2242
Qy	715	ValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaIle	734
Db	2243	TTGGTCCGATGTGACACACCCACCAAGTGGCGGACAAACAAGAAACCATCGATTGCGCC	2302
Qy	735	ValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGlu	754
Db	2303	GTGCTGGCTCCATGATGCCATCCATCGCGTATGCGCCACCGTTCCGCTACAGCAG	2362
Qy	755	GluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu	774
Db	2363	-----CACCGACAGGAGATCATCCAGGAGCTGAGC	2392
Qy	775	AlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAla	794
Db	2393	AGCATGTGCGCGAGCTGTTGATCATGTTCTACAAGTGCAGCGGGGCTACAAGCCCCAC	2452
Qy	795	HisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAsp	814
Db	2453	CGCATCATCTCATCGTGAAGAGTCTCCGAGGACAATCCACATGCTCTCTGCAACAC	2512
Qy	815	GluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAsp	834
Db	2513	GAATTGACCCCATTCGCGAGGCGCTGCATTAGCTA-----GAG	2551
Qy	835	ProGlu-----ProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeu	852
Db	2552	CCAGAGTATCGCGCGGCATCATATTGTTGTGCAGAGAGCGCCATCATCACTGACTC	2611
Qy	853	LeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThr	872
Db	2612	TTCTGCGCGAGAGAAGGAG-----	2632
Qy	873	AspValAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThr	892
Db	2633	-----CAGAGCGGCAATCG	2647
Qy	893	GlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLys	912

```

Db      2648 GGC-----ATATTCGCGAGCACCACCGTCGATGGGCATCACACATCCACCGAA 2701
Qy      913 PheAspPheLeuAlaSerHisHisGlyValLeuGlyThrSerArgProGlyHisTyr 932
Db      2702 TTTGATTCTCTATCTATGCGCCATCAGGGCATCCAGGGCACCAGTCGCCCTCGCACTAC 2761
Qy      933 ThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGly 952
Db      2762 CACGTTCTGTGGACGACAATCACTTTGACTCGGACGAGCTGCAGTGCCTCACGTATCAG 2821
Qy      953 LeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyr 972
Db      2822 CTATGCCATACGTACGTGGCTGCACCCGATCCGTCACTATACCGCGCGCCAGCTACTAC 2881
Qy      973 AlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIle 992
Db      2882 GCCCATTTAGTGGCCTTCGGTGCAGA--TATCATCTGGTGAGAGAAGGAGCAC----- 2932
Qy      993 GlyAsp-TyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnThrAsnVa 1012
Db      2933 ---GATTCGGCGGAGGTTTCGCACACAGAGCGGCTGCTCAGAGGATCGTACGC---CAGGT 2986
Qy      1012 lLysTyrProGlyMetSer 1018
Db      2987 GCCATGGCCAGGGCCATCA 3005

```

Search completed: July 6, 2006, 02:32:01
 Job time : 2800 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 6, 2006, 01:29:24 ; Search time 1437 Seconds
(without alignments)
1384.275 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 5349

Sequence: 1 MSNPFELKXGFRHSLDFE.....RHEHFLQTNVKYQMSFA 1020

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abes/ABSSWEB_spool/US10645746/runat_05072006_110200_5345/app_query.fasta.1
-DB=Published Applications NA.New -QWMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANSHUMAN40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abes03p
-USER=US10645746 @CEN 1.1.206 @runat_05072006_110200_5345 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /EMC Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	577	10.8	7478	US-11-266-748A-56189	Sequence 56189, A
2	555	10.4	4758	US-11-266-748A-30266	Sequence 30266, A
3	535	10.0	3474	US-10-953-349-4063	Sequence 4063, Ap
4	510	9.5	3320	US-11-266-748A-31508	Sequence 31508, A
5	510	9.5	3320	US-11-266-748A-56999	Sequence 56999, A
6	509	9.5	3414	US-10-449-902-24843	Sequence 24843, A
7	500	9.3	3432	US-10-449-902-27837	Sequence 27837, A
8	499.5	9.3	3746	US-11-218-305-22486	Sequence 22486, A

9	493	9.2	1793	8	US-11-266-748A-25558	Sequence 25558, A
10	486	9.1	2503	6	US-10-449-902-19320	Sequence 19320, A
11	479	9.0	5408	7	US-11-218-305-543	Sequence 543, App
12	469.5	8.8	3350	7	US-11-218-305-19388	Sequence 19388, A
13	466	8.7	2012	7	US-11-218-305-14683	Sequence 14683, A
14	456	8.5	2773	7	US-11-218-305-19467	Sequence 19467, A
15	455	8.5	3827	7	US-11-218-305-19466	Sequence 19466, A
16	453.5	8.5	2808	8	US-11-216-545-2985	Sequence 2985, Ap
17	443.5	8.3	3178	6	US-10-449-902-14723	Sequence 14723, A
18	431.5	8.1	3003	6	US-10-449-902-8899	Sequence 8899, Ap
19	421.5	7.9	1750	6	US-10-449-902-10173	Sequence 10173, A
20	420.5	7.9	3047	6	US-10-449-902-14194	Sequence 14194, A
21	412.5	7.7	2016	7	US-11-218-305-22484	Sequence 22484, A
22	403.5	7.5	1754	6	US-10-953-349-18472	Sequence 18472, A
23	397	7.4	2001	7	US-11-218-305-22483	Sequence 22483, A
24	396.5	7.4	1816	6	US-10-953-349-19397	Sequence 19397, A
25	392.5	7.3	2095	7	US-11-218-305-1556	Sequence 1556, Ap
26	377	7.0	1710	7	US-11-218-305-19385	Sequence 19385, A
27	374	7.0	3139	7	US-11-218-305-20989	Sequence 20989, A
28	358.5	6.7	1934	8	US-11-216-545-6852	Sequence 6852, Ap
29	344.5	6.4	1211	6	US-10-449-902-26400	Sequence 26400, A
30	340	6.4	1324	8	US-11-216-545-2982	Sequence 2982, Ap
31	331	6.2	1196	6	US-10-449-902-3167	Sequence 3167, Ap
32	331	6.2	1196	6	US-10-449-902-3806	Sequence 3806, Ap
33	331	6.2	1298	8	US-11-216-545-2984	Sequence 2984, Ap
34	312	5.8	3394	8	US-11-266-748A-26618	Sequence 26618, A
35	312	5.8	3394	8	US-11-266-748A-29538	Sequence 29538, A
36	286.5	5.4	1094	7	US-11-218-305-21995	Sequence 21995, A
37	283.5	5.3	965	7	US-11-218-305-22485	Sequence 22485, A
38	260.5	4.9	2456	8	US-11-293-697-1525	Sequence 1525, Ap
39	259	4.8	4753	7	US-11-218-305-19384	Sequence 19384, A
40	258.5	4.8	892	7	US-11-218-305-19386	Sequence 19386, A
41	253.5	4.7	997	8	US-11-216-545-6851	Sequence 6851, Ap
42	249	4.7	2997	6	US-10-449-902-27469	Sequence 27469, A
43	232.5	4.3	772	6	US-10-449-902-4254	Sequence 4254, Ap
44	232	4.3	903	7	US-11-218-305-19468	Sequence 19468, A
45	206	3.9	1000	8	US-11-266-748A-394325	Sequence 394325, A

ALIGNMENTS

RESULT 1
US-11-266-748A-56189
; Sequence 56189, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56189

491 GATAGGTTGATATGAGGTGACTCTTCCAGCGGGTAAAGACCAACATTTAAAGTG 550
155 -----IleLeuTyrThrMetIleLeuTyrThrArgLysPheHis 168
551 TCTGTTTCAGTGGGTGTCAGTTGTGTGAGCCCTTCAGTTGCTTTTAAAGCTTTGGCTGGCAC 610
169 LeuAenPheSerArgGluAenProGluLysAenGluGluAaAenArgSerTyrLysPhe 188
611 TTGAAT-----GAATCCAGATGACTCAGTACAAGCA----- 643
189 LeuLysAenValMetThrGln-----LysValArgTyrAlaProPheValAenGlu 205
644 ---CTTGATGTTATCACAAGACACCTTCCCTCCATGAGGTACACCCAGTG----- 691
206 GluIleLysValGlnPheAlaLysAenPheValTyrAspAenSerIleLeuArgVal 225
692 -----GGCCGTTCCTTTTCTCACCC 712
226 ProGluSerPheHisAspProAenArgPheGluGlnSerLeuGluValAlaProArgIle 245
713 CCGGAGGTACTACACCT-----CTGGAGGGGGCAGG 748
246 GluAlaTrpPheGlyIleTyrIleGlyIleLysGluLeuPheAspGlyGluProValLeu 265
749 GAGGTCTGCTGTTGTTTTCATCAGTCTGTGAGACCTGCCATG---TGGAAATATGATGCTC 805
266 AenPheAlaIleValAspLysLeuPheTyrAenAlaProLysMetSerLeuLeuAspTyr 285
806 AACATTGATGTATCTGCAACTCTTCTACCGGCT-----CAGCCTATCATTTAGTTTC 859
286 LeuLeuLeuIleValAspProGlnSerCysAenAspValAlaArgLysAspLeuLysThr 305
860 ATGTGTGAGTGTATAGACATTTCAGAACATCATGTAAACACCAAA----- 904
306 LysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeu 325
905 -----CCTCTAACAGACTCCAGCGGTGTCAAATTTACCAAGAA 943
326 LeuGluAenLeuLysLeuLys-----CysAlaGluValTrpAspAen----- 339
944 ATCAGAGGTCTCAAAAGTTGAGGTGACCCACTGTGGACAGATGAACGAAATACCGAGTT 1003
340 ---GluMetSerArgLeuThrGluArgHisLeuThrPhe---LeuAspLeuCysGluGlu 357
1004 TGTAATGTACGTAGACGGCCAGCAGTCAATCAACTTTCCTTTGTCAGCTA----- 1054
358 AenSerLeuValTyrLysValThrGlyLysSerAspArgGlyArgAenAlaLysLysTyr 377
1055 -----GAAAAACGGTCAAGCTATG 1072
378 AspThrThrLeuPheLysIleTyrGluGluAenLysLysPhe---IleGluPheProHis 396
1073 GAATGTACAGTAGCTCAATATTTTAAAGCAAAAGTATAGTCTGCAACTGAAATACCCCAT 1132
397 LeuProLeuValLysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeu 416
1133 CTTCCTGTCTCCAGTGGGACAGNACAAGACATACATACTGCCACTCGAGGTCTGT 1192
417 GluValHisGluLysProGlnArgTyrLysAenArgIleAspLeuValMetGlnAspLys 436
1193 AATATA---GTGGCAGGACAGCGATGTATCAAGAAAGCTCACAGACAAATCAGACTTCCACA 1249
437 PheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAenThrLeuLysMetLeu 456
1250 ATGATCAAAAGCTACAGCAAGATCTGCTCTGTGACAGACAGNAGAGATCAGTAGAGTGGTG 1309
457 LysGluLeuAspPheSerSerGluGluLeuAenPheValGluArgPheGlyLeuCysSer 476
1310 AAGACCAACAGTATGTTGGGTGGACCTGTATCCTGTGACAGACAGNAGAGATCAGTAGAGTGGTG 1369
477 LysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeuValAenSer 496
1370 CACAATGAAATGACAGACTCACAGCAGGGTACTTCCAGCACCACCAATGCTGCAATATGGA 1429

497 ValAenGluGlnIleLysMetThrPro-----ValIleArgGlyPheGln 511
1430 GGCCGGGAATAAACAGTAGTACCCACCCAAACCCAGGGTGTCTGGAGATCGCAGGA----- 1483
512 GluLysGlnLeuAenVal---ValProGluLysGluLeuCysCysAlaValPheValVal 530
1484 ---AAGCAGTTTATGCTGGCATTGAAATTAAGTTTGGGCAGTGTCTTTTGCACCT 1540
531 AenGluThrAlaGlyAsnProCysLeuGluGluAenAspValValLysPheTyrThrGlu 550
1541 CAGAAA-----CAATGTAGGAA-----GATTTACTAAGAGATTTCACTGAC 1582
551 LeuIle-----GlyGlyCysLysPheArgGlyIleArgIleGlyAla 564
1583 CAGCTGCGTAAATCTTAAGATGCAAGATGCCATCCAGGGTCAGGCATGTTTCTGTC 1642
565 AenGluAenArgGlyValaGlnSerIleMetTyrAspAlaThrLysAenGluTyrAlaPhe 584
1643 AAGTATGCACAAGGTGCAGACAGTGTG----- 1669
585 TyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaThrGluAla 604
1669 ----- 1669
605 LysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIleIle 624
1670 GAGCCTATGTTTAAACATCTGAAATAGCTATATGTGGCCTACAGCTAATAGTGTATC 1729
625 SerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHisThrIleGly 644
1730 CTGCTCGAAAGACACACCATATATCGGAGGTGAAACGTTGTGGAGATACCTTCTAGGT 1789
645 ValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHis 664
1790 ATGGCCACACAGTGTGTCCAGGTAAAAATGTAGTGAAGACC----- 1831
665 GluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAenAlaLysLeuGly 684
1832 -----TCACCTCAAAACCTTCCAAATCTTTCCTGGAAGTAAATGCAAACTTGA 1882
685 GlyIleAenGlnLeuAspTrpSerGluIleAlaGluIleSerProGluGluLysGlu 704
1883 GGAATTAACAATGTGCTT-----GTGCCTCATCAAGGCC 1918
705 ArgArgLysThrMetProLeuThrMetTyrValGlyIleAspValThrHisProThrSer 724
1919 TCGGTGTTCCAGCACCTGTC---ATCTTCTGGAGCGGATGTACACACCCCCAGCA 1975
725 TyrSerGlyIleAspTyrSerIleAlaValAlaSerIleAenProGlyGlyThr 744
1976 GGGGATGGAGAAACCTTCCATTCCTGCTGTGGTGGCAGTATGATGCCACCCACG 2035
745 IleTyrArgAenMetIleValThrGlnGluCysArgProGlyGluArgAlaValAla 764
2036 CGGTACTGTCCACCGTTCGGGTGCAGACTTCCGCGCAGGAGATCTCCCAAGAGCTCCTC 2095
765 HisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPhe 784
2096 TACAGTCAAGAGGTGCATCCAGACCTGACTAAACATGGTTCGAGAGCTGCTGATTCAGTTC 2155
785 AlaGluAenAenAspAenArgAlaProAlaHisIleValValTyrArgAspGlyValSer 804
2156 TACAATCCACAGCTTCAAA---CCACTCGATCATCTATACCGTGGGGGTATCT 2212
805 AspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLys 824
2213 GAGGACAAAATAAGACAGTGTGGCCAGAACTAATAGCAATTCGA-----AAG 2263
825 GlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIle 844
2264 GCATGTATTAGTTCGAA-----GAAGATTACCGCCAGGAATAACTTATTATTGGTGT 2317

```
QY 845 GlnLysArgHisAsnThrArgLeuArgMetGluLysAspLysProValValAsn 864
Db 2318 CAAAAAGACATACACAGACTCTTC---TGTCAGATAAACAGAAAGGGTAGGAAA 2374
QY 865 LysAspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGlu 884
Db 2375 AGTGGCAATGTACCAGCA----- 2392
QY 885 AspMetLysGluSerLysGluThrGlyLeValAsnProSerSerGlyThrValAsp 904
Db 2393 -----GGCACTACAGTGGAT 2407
QY 905 LysLeuLeValSerLysTyLysPheAspPheLeuAlaSerHisHisGlyValLeu 924
Db 2408 AGTACCATCACACATCATCTGAGTTGACTTTTACTCTGTAGTCATGAGGAATTCAG 2467
QY 925 GlyThrSerArgProGlyHisTyThrValMetTyLysAspLysGlyMetSerGlnAsp 944
Db 2468 GGAACACAGCGTCCCTCACATTACAGGTCTTGTGGGATGACAACTGCTTTCAGTCAGAT 2527
QY 945 GluValTyLysMetThrTyThrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIle 964
Db 2528 GAACTCCAGCTACTGACTTACCAGCTGTGTACACCTATATGTAGGTGCACTCGCTCAGTC 2587
QY 965 SerLeuProValHisTyAlaHisLeuSerCysGluLysAlaLysGluLeuTyLys 984
Db 2588 TCTATCCAGCCCTGCATATATATGCCGCTTGTAGCATTAGGGCAAGG---TATCAT 2644
QY 985 ArgThrTyLysGluHis 990
Db 2645 CTGGTGGATAAAGATCAT 2662
```

RESULT 3

```
US-10-953-349-4063
; Sequence 4063, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4063
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-4063
```

```
Alignment Scores:
Pred. No.: 2,32e-44 Length: 3474
Score: 535.00 Matches: 231
Percent Similarity: 38.8% Conservative: 184
Best Local Similarity: 21.6% Mismatches: 370
Query Match: 10.0% Indels: 284
Db: Gaps: 42
```

US-10-645-746-3 (1-1020) x US-10-953-349-4063 (1-3474)

```
QY 1 MetSerSerAsnPheProGluLeu-----GluLysGlyPheTyArgHisSerLeuAsp 18
Db 586 CTGGCTCAGCAATTTGAACAACACTCTCTGTGTGAACAAGAGAGCTCCCAAGTCAGGCAATCCAG 645
QY 19 Pro-----GluMetLysTrpLeuAlaArgProThrGlyLysCysAsp 32
Db 646 CCTATACCTTCTCTAGCAGGCTTCAAGTTTCCATGAGGCTCGTAAAGACAGAGT 705
QY 33 GlyLysPheTyGluLysLysValLeuLeuValAsnTrpPheLysPheSerSerLys 52
Db 706 GGA-----AAGGTTGCATTGTGAAGGCTAAC-----CATTTCTTTGCTGAA 747
```

```
QY 53 IleTyArgAspArgGluTyTyTyGluTyGluValLysMetThrLysGluValLeuAsnArg 72
Db 748 CTCCTCTAAGGATTTGGACCATATATGATGTTTACCATTACTTCGGAAGTTTACATCAAGG 807
QY 73 LysProGlyLysProPheProLysLysThrGluIleProIleProAspArgAlaLysLeu 92
Db 808 GGTGTCAATCGTCTGTGATGAACAA----- 834
QY 93 PheTrpGlnHisLeuArgHisGluLysGlnThrAspPheIleLeuGluAspTyLys 111
Db 835 -----CTTGTGATAATATCGT 852
QY 112 -----ValPheAspGluLysAspThrValTyLysSerVal 122
Db 853 GATTCTCACCTTGGAAAGTCGTCTCCAGCGTATATGTCGAAAGACTCTTTACACTGCT 912
QY 123 CysArgLeu-----AsnThrValThrSerLysMetLeuValSerGluLysValValLys 140
Db 913 GGTCCACTTCCCTTTAACTCCAAGGAGTTCAGAAATCAATCTTCTTGGACGAAGAGTGGG 972
QY 141 LysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeuTyThrMetIle 160
Db 973 GCTGGAGGTCAAAGACGAGAAAGG-----GAATTTAAAGTTGTGATCAAGCTAGTT 1023
QY 161 LeuThrTyArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAsp 179
Db 1024 ---GCACGTGCTGATCTGCATCACCATTAGGAATGTTTTTGGAGGGGAAACAATCAGATGCC 1080
QY 180 ---GluGluAlaAsnArgSerTyLysPhe-----LeuLysAsnValMetThrGlnLysVal 197
Db 1081 CCACAGAGAGCTCTGCAGGTTCTTGACATTTGTTCTGTCGAGCTGCCGACCTCTAGATC 1140
QY 198 ArgTyAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTyLys 217
Db 1141 AGGTATATTCG----- 1152
QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
Db 1153 -----GTGGCCCGTCTCTTTTATTTCCCTCGAT---ATAGGAAAA 1188
QY 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyLysIleGlyLys 256
Db 1189 AAACATCATTTGGGGATGCTTGGAGAGCTGGCTGGATTTCTACCAAGCATTCGTCCT 1248
QY 257 -----GluLeuPheAspGlyGlu 262
Db 1249 ACACAGATGGGCTTATCACTCAATATTGATATGTCATCGACAGCTTCATAGAGCAAC 1308
QY 263 ProValLeuAsnPheAlaIleValAspLysLeuPheTyAsnAlaProLysMetSerLeu 282
Db 1309 CCTGTGATTTCAGTTT-----GTC 1326
QY 283 LeuAspTyLysLeuLeuIleValAspProGlnSerCysAsnAspAspValAlaGlyAsp 302
Db 1327 TGTGATTTCTTAACCGGGATATTTCTCTCGACCTTTATCTGATGCTGATCTGTTAAG 1386
QY 303 LeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgProArgIle 322
Db 1387 ATAAAAAAGGCTCTTAGAGGTGTCAAAGTTGAAGTCACTCATCGAGAAACATCGCCCG 1446
QY 323 ArgGlnLeuLeuGluAsnLeuLysLysCysAlaGluValTrpAspAsnGluMetSer 342
Db 1447 AAGTACCGCATTTCCGGTTTG-----ACT 1470
QY 343 ArgLeuThrGluArgHisLeuThrPhe-----LeuAspLeuCysGluGluAsnSerLeuVal 361
Db 1471 GCTGTGGCCACTCGGAATTGACATTCACAGTAGAT----- 1506
QY 362 TyLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyAspThrThrLeu 381
Db 1507 -----GAAAGAAATACTCAGAAA-----TCTGTT 1530
QY 382 PheLysIleTyGluGluAsnLysLysPhe-----IleGluPheProHisLeuProLeuVal 400
```

Db	1531	GTAGAACTTCCACGAAACATATGGTTTGGCATTCAGCACACTCACTACCATCGTGTG	1590
Qy	401	LysVallylSerGlyAlaLysGluTyAlaValProMetGluHisLeuGluValHisGlu	420
Db	1591	CAAGTTGGGAACTTAATAGCCCTAAATTACTTACCAATGGAGGTATGCAAGATGTTGAA	1650
Qy	421	LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArg	440
Db	1651	--GGCCAGCGGTATTCCAAAGATTGAATGAGAGACAGATCACTGCTTGTCTGAAGGTT	1707
Qy	441	AlaThrArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAsp	460
Db	1708	ACCTGTGAGCGCCGATAGATCGAGAAAAAGATATCTTACAGACGGTGCACCACTCAATGAT	1767
Qy	461	PheSerSerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMet	480
Db	1768	TATGCTAAAGAT-----AATTATGCTCAAGAGTTTGGCATCAAAATAAGTACTTCTCTG	1821
Qy	481	IleGluCysProGlyLysValLeuLysGluProMetLeu--ValAsnSerValAsnGlu	499
Db	1822	GCTTCTGTGAGGCTCGTACTACTCCTCTCCATGGCTTAAGTACCACGAGCTGGAAGG	1881
Qy	500	GlnIleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValPro	519
Db	1882	GAAGGCACTGTCTGCCA-----CAAGTTGCTCAATGGACATGATGAAT	1926
Qy	520	GluLysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsn--ProCys	538
Db	1927	AAGAAA-----ATCATCAATGGTGAACGGTGAATATATGGATCTGC	1968
Qy	539	Leu-----GluGluAsnAspValValLysPheTyrThrGluLeuIle	552
Db	1969	ATCAACTTTTCTAGGCAAGTCAGGACACATCTACCGGTACATTTTGTCTCAGAACTTGCT	2028
Qy	553	GlyGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSer	572
Db	2029	CAAAATGTTTACGTATCTGGCATG-----	2052
Qy	573	IleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThr	592
Db	2053	-----GCATTTAATCCGGAACAGTCCTCCCAACCA	2082
Qy	593	GlyIleGlyArgPheGlu-----IleAlaAlaThrGluAlaLysAsnMetPhe	608
Db	2083	GTCAGTCTCGCCTGAGCAAGTAGAGAAAGGTCTTGAAGACTAGATATCATGATGCCACA	2142
Qy	609	GluArgLeuProAspLysGluGlnLysValLeuMetPheIleIleSerLysArgGln	628
Db	2143	TCAAAACTCTCCCAAGGAAAGAAATATGATCTGCTTATGTCATTCTGCCGATAATAAT	2202
Qy	629	LeuAsnAlaTyrGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGln	648
Db	2203	GGATCATTTATACGGTGATTTGAAACGCATATGTGAGACTGAACCTGGCATGTCTCTCAA	2262
Qy	649	HisIleThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySer	668
Db	2263	TGTTGCTGTACAAAGCATGCTTTAAGATG-----AGC	2295
Qy	669	LysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGln	688
Db	2296	AAACAATACATCGGCTAATGTTGGCTGAAGATTAATGTGAAGGTTGCGAAGAAACACA	2355
Qy	689	GluLeuAspTyrPsrGluIleAlaGluIleSerProGluGluLysGluArgArgLysThr	708
Db	2356	GTGCTT-----GTTGATGCTCTATCT-----AGCGG	2382
Qy	709	MetProLeu-----ThrMetTyrValGlyIleAspValThrHisProThr	723
Db	2383	ATTCCTCTAGTCACTGATCGACCCACCATTAATTTGGTGTGATGATTCACCCCTCAC	2442
Qy	724	SerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlaSerIleAsn--ProGly	742

2443 CCTGGAGAGATTCAAGCCCATCTATTGCTGCTGTTGTGTGGCATCTCAGGATTGGCCCTGAA 2507
743 GlyThrIleTyrArgAsnMetIleValThrGln----- 753
2503 ATCACTAAATATGCTGGATTAGTTTGGCTCAAGCGCATAGGAGGAGCTCATTCAGGAT 2562
754 -----GluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluAthrAep 771
2563 CTGTTTCAAGAGTGAAGGATCTCTCAGAAAGGTGTGTGACTGGT----- 2607
772 IleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAepAepAArg 791
2608 -----GGCATGATAAAGGAGTTGCTCATAGCCTTCCTGATAGATCAACTGGCGCATAA 2658
792 AlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgVal 811
2659 ---CCACTAAGGATCATCTTCTACAGGATGGAGTCAGTGAGGACAAATTTTACCAAGTT 2715
812 SerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluAthrAep 831
2716 TTGCTCTATGAACCTTGATGCCATCCGC-----AAGGCCTGTGCTCGCTGGAGCA 2766
832 GlyClnAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArg 851
2767 GGTAT-----CAACACACAGTGACATTTGTGGTGTGCAGAAAGCGTCATCACAGAGG 2820
852 LeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGlu 871
2821 CTGTTTCTCAGAACCAACATGATGCCATTCGGTGGACAGA----- 2862
872 ThrAspValAlaValAlaValLysGlnTrpGluAspMetLysGluSerLysGlu 891
2862 ----- 2862
892 ThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyr 911
2863 AGTGGGAATATTTTACCT-----GGCAGCTGTTGTGGACTCTTAAATCTGCCACCTTACA 2916
912 LysPheAspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHis 931
2917 GAGTTTGACTTTTACCTCTGTAGTCATGCTGTGATTTCAGGCGCACTTCTGACCTGCTCAT 2976
932 TyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyr 951
2977 TACCACGTTCTTTGGGATGAGAACACTTTTACTGCAGATGGACTTCAATCTCTGACCAAT 3036
952 GlyLeuAlaPheLeuSerAlaAArgCysArgLysPheProIleSerLeuProValProValHis 971
3037 AACTTATGTTACACGATGCAAGATGCACACGCTCAGTTTCAANTGTTCCCTCGCATAT 3096
972 TyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyr 991
3097 TATGCACATCTAGCAGCTTTTAGGGCTCGA----- 3126
992 IleGlyAspTyrAlaGlnProArgThr 1000
3127 -----TTCTACATGGAGCCAGACAGACA 3147

RESULT 4
US-11-266-748A-31508
; Sequence 31508, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIORITY FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIORITY FILING DATE: 2004-11-03

```
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31508
; LENGTH: 3320
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-31508

Alignment Scores:
Pred. No.: 7,77e-42 Length: 3320
Score: 510.00 Matches: 189
Percent Similarity: 42.8% Conservative: 127
Best Local Similarity: 25.6% Mismatches: 264
Query Match: 9.5% Indels: 160
DB: 26 Gaps: 26

US-10-645-746-3 (1-1020) x US-11-266-748A-31508 (1-3320)

QY 274 PheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuLeuLeuValAspProGln 293
DB 671 TTCTACAAAGCA-----CAACCTGTAATTCAGTTGATGTTGAGTTTCTGATATTCAT 724
QY 294 SerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThr 313
DB 725 AATATTGATGAGCAACCAAG-----CCT 748
QY 314 IleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuLeuLeuLeuLysLysCys 333
DB 749 CTGACTGATTTCTCATCGGGTAAATTCACCAAGAGATAAAGGTTTGAAGTTTGAAGTG 808
QY 334 AlaGluValTyrAspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeu--- 352
DB 809 AC-TCATTGTGGAACAATGAG-----ACGGAA-----ATACCGTGTGTTGTAA 849
QY 353 -----AspLeuCysGluGluAenSerLeuValTyrLysValThrGlyLysSerAsp 369
DB 850 TGTAACAAGAGCGCTGCGCAGTCATCAACCTTTTCCTTTACAGTTAGAAACGCGCAAC 909
QY 370 ArgGlyArgAsnAlaLys-LysTyrAspThrThrLeuPheLysIleTyrGluGluLeuLys 389
DB 910 TGTGGAGAGAACAGTAGCGCAGTAT-----TTTCAGAGAAAGTATAC 951
QY 389 slyPheIleGluPheProHisLeuValLysValLysSerGlyAlaLysGluTy 409
DB 952 TCTTCAGCTGAAGTACCGCACCCTTCCTGTCGCAAGTGGCGCAGAAACAGAAACACAC 1011
QY 409 xAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArg11 429
DB 1012 CTACCTGGCCACTAGAAGTCTGTAATATT---GTGGCAGGCGCAACGATGTATCAAGAGCT 1068
QY 429 eAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLy 449
DB 1069 AACAGACATCAGACTTCCCATATGATCAAGGCAACAGCAGATCTGCACCAGATAGACA 1128
QY 449 sGluAenThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVa 469
DB 1129 AGAGGAAATATGACAGATGTAAGAAGTGCMAATTATTAACACAGATCCA-----TTTGT 1182
QY 469 lGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLy 489
```

```
Db 1183 TCAGGAGTTTCAATTTAAAGTTCGGGATGAATGGCTCATGTAACTGGACGCTACTTCC 1242
QY 489 sGluProMetLeuValAenSerValAenGluGlnIleLysMetThrProVal----- 506
Db 1243 AGCACCTATGCTCCAGTATGGAGCAGGAATCGGACAGTAGTACACACCGACCCATGGAGT 1302
QY 507 -----IleArgGlyPheGlnGluLysGlnLeuAenVal---ValProGluLysGluLe 523
Db 1303 ATGGGACATCGAGGG-----AAACAATTCCACACAGGAGTTGAAATCAAAATGTG 1353
QY 523 uCysCysAlaValPheValValAenGluThrAlaGlyAsnProCysLeuGluGluAenAs 543
Db 1354 GGCATTCGCTTGTGTTTTCACACAG-----AGGCAGTGCAGAGAGAA----- 1396
QY 543 pValValLysPheTyrThrGluLeuLe-----GlyGlyCysLysPhe 557
Db 1397 -ATATTGAAGGGTTTCACACAGCAGCTGCGTAAGATTTCTAAGATGCGGGATGCCCAT 1455
QY 557 eArgGlyIleArgIleGlyAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAl 577
Db 1456 CCAGGCGCAGCCATGCTTCTGCAAAATATGACAGGGGGCAGACAGCGTA----- 1504
QY 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAenThrGlyIleGlyArgPh 597
Db 1504 ----- 1504
QY 597 eGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLy 617
Db 1505 -----GAGCCCATGTTCCGGCATCTCAAGAACACATATTCTCGG 1542
QY 617 sValLeuMetPheIleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHi 637
Db 1543 CCTACAGCTATTATGTCATCTCGCGGGAACACACAGTAGTGTATCGGAAGTGAACG 1602
QY 637 sTyrCysAspHisThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLy 657
Db 1603 TGTAGGAGACACACTTTTGGGTATGGCTACACAATGTGTTCAAGTCAAGATGTGAATAA 1662
QY 657 sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLe 677
Db 1663 AACA-----TCTCTCAAACTCTGCAAACTGTGCCT 1695
QY 677 uLysIleAenAlaLysLeuGlyGlyIleAenGlnLeuLeuAspTyrSerGluIleAlaGl 697
Db 1696 AAGATAAATGTTAAACTCGGAGGATCAATAATATCTT----- 1735
QY 697 uIleSerProGluLysGluArgLysThrMetProLeuThrMetTyrValGlyI1 717
Db 1736 ----GTACCTCATCAAGAGACCTTCTGTGTTCCAGCAACCAAGTG---ATCTTTTGGGAGC 1788
QY 717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAl 737
Db 1789 CGATGTCTCATCTCCACCTGCTGCTGATGGAAGAAGCCCTTCTATTGCTGCTGTTGAGG 1848
QY 737 aSerIleAenProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluGluCysAr 757
Db 1849 TAGTATGGATGCACACCCCAAGCAGATCTGTGCCACAGTAAGAGTTTCAG----- 1897
QY 757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
Db 1898 -----AGACCCCGCAGGAGATCATCCAGACTTGGC 1929
QY 775 aLysPheValLys---LeuLeuArgGluPheAlaGluAenAsnAspAenAlaProAl 794
Db 1930 CTCCATGTCGGGAACCTTCTATTCAATTTTATAAGTCAACTCGGTTCAAG---CCTAC 1986
QY 794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
Db 1987 TCGTATCATCTTTATCGGGATGGTGTCTTTCAGAGGGGCGAGTTTTCAGGAGTATATATTA 2046
QY 814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGlu-A 834
```

```
Db 2047 TGAACCTACTAGCAATTCGAGA-AGCGTCGCATCAGTTT-----GGAGAAAG 2090
Qy 834 sPProGluProLysTyrThrPheleVallelGlnLysArgHisAsnThrArgLeuLeuA 854
Db 2091 ACTATCAACCTGGAATTAACCTACATTTGATTTGAGAAGAGACATCACACTCGATTATTT- 2149
Qy 854 rGArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspV 874
Db 2150 ---TGTGCTGATAGGACAGAAAGGTTGGAAAGAGTGGCAATATCCAGCT----- 2197
Qy 874 alalaValAlaAlaValLysGlnTrpGluLysAspMetLysGluSerLysGluThrGlyI 894
Db 2197 ----- 2197
Qy 894 leValAsnProSerSerGlyThrThrValAspLysLeuLeuValSerLysTyrLysPheA 914
Db 2198 -----GGAACAACAGTTTGATACAGACATTAACACCCCATATGAGTTGCG 2240
Qy 914 sPhePheLeuAlaSerHisGleGlyValLeuGlyThrSerArgProGlyHisTyrThrV 934
Db 2241 ATTTTACCTCTGTGACCACTGCTGGAATACAGGGTACCGAGTCGCTCCTTCACACTATCATG 2300
Qy 934 alMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuA 954
Db 2301 TTTTATGGATGATACTGCTTTACTGCGAGATGAACCTTCAGCTGCTTAACCGCTCT 2360
Qy 954 laPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaH 974
Db 2361 GCCACACTTACGTACGCTGTACACGATCTGTTTCTATACCTGCACCGCGGTATTATGCTC 2420
Qy 974 laLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluLys 990
Db 2421 ACCTGTGATGATTTAGAGCCAGA---TATCATCTTGTGGACAAGAAACAT 2467
```

RESULT 5

```
US-11-266-748A-56999
; Sequence 56999, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56999
; LENGTH: 3320
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-56999
```

Alignment Scores:

Pred. No.: 7.77e-42 Length: 3320

```
Score: 510.00 Matches: 189
Percent Similarity: 42.8% Conservative: 127
Best Local Similarity: 25.6% Mismatches: 264
Query Match: 9.5% Indels: 160
DB: 8 Gaps: 26

US-10-645-746-3 (1-1020) x US-11-266-748A-56999 (1-3320)
Qy 274 PheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuLeuValAspProGln 293
Db 671 TTCTACAAGCA-----CAACCTGTGAATTCAGTTTCATGTGTGAAGTTCTTGATATTCAT 724
Qy 294 SerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThr 313
Db 725 ATATTGATGAGCAACCAAGA-----CCT 748
Qy 314 IleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLysCys 333
Db 749 CTGACTGATTCTCATCGGTAATAAATTCACCAAGAGATAAAAGGTTTGAAGGTTGAAGTG 808
Qy 334 AlaGluValTrpAspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeu--- 352
Db 809 AC-TCAATTGTGGAACAATGAG-----ACGGAA-----ATACCGTGTGTTGTA 849
Qy 353 -----AspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAsp 369
Db 850 TGTAAACAGGAGCGCTGCAGTCATCAAACTTCTTACAGTTAGAAAACGGCCAAAC 909
Qy 370 ArgGlyArgAsnAlaLys-LysTyrAspThrThrLeuPheLysIleTyrGluGluAsnly 389
Db 910 TGTGAGAGAACAGTAGCGCAGTAT-----TTCAGAGAAAAGATATAC 951
Qy 389 sLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTy 409
Db 952 TCTTCAGCTGAAGTACCGCCACCTTCCCTGCTGCAAGTCGGCGAGGAACAGAAAACACAC 1011
Qy 409 rAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgI 429
Db 1012 CTACCTGCCACTAGAAAGTCTGTAATATT---GTGCGAGGCAACCGATGTTATCAAGAAGCT 1068
Qy 429 eAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLy 449
Db 1069 AACAGACAATCAGACTTCCACTATCATCAAGGCAACAGCAAGATCTGCACCAGATAGACA 1128
Qy 449 sGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVa 469
Db 1129 AGAGGAAATATGACAGATTGGTAAGAAAGTGCAAAATTATGAACAGATCCA-----TTTGT 1182
Qy 469 lGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLy 489
Db 1183 TCAGGAGTTTCAATTTAAAGTTCCGGATGAAATGGCTCATGTAACTGGACGGTACTTCC 1242
Qy 489 sGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal----- 506
Db 1243 AGCACCTATGCTCCAGTATGGAGGACGGAATCGGACAGTAGCAACACCCAGGCATGAGCT 1302
Qy 507 -----IleArgGlyPheGlnGluLysGlnLeuAsnVal---ValProGluLysGluLe 523
Db 1303 ATGGGACATGCGAGGG-----AAACAATTCCACAGAGGTTGAAATCAAAATGTG 1353
Qy 523 uCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAs 543
Db 1354 GGCTATCGCTGTTTTCACACAG-----AGGCAGTGCAGAGAGAA----- 1396
Qy 543 pValValLysPheTyrThrGluLeuLe-----GlyGlyCysLysPh 557
Db 1397 -ATATTGAAGGGTTTTCACAGACCGCTCGCTGAAGATTTCTAAGGATGCGAGGATGCCCAT 1455
Qy 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
Db 1456 CCAGGCGCCGCCCATGCTCTCTGCAAAATATGCACAGGGGGGCGACAGCGGTA----- 1504
Qy 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPh 597
```


Db	922	ATGGACTTCAGATCAAGCTGGCAAGAGCTCAGGGAATCGAGTACAGGTG-----	795	Db	1816	GGGCCCAACGCTG-----	1830
Qy	158	ThrMetIleLeuThrTyrArgLysLys-----PheHisLeuAenPheSer	172	Qy	510	PheGlnGluLysGlnLeuAenValValProGluLysGluLeuCysCysAlaValPheVal	529
Db	976	-----ACAAATTCGACGTGCTGGCCAGCAGATCTGCACCACTGCACGAGTTC	1023	Db	1831	TACCATGACTCT-----GGAAAGGAGAAAATTGT	1860
Qy	173	ArgGluAenProGluLysAspGluAlaAenArgSerTyrLysPheLeuLysAenVal	192	Qy	530	ValAenGluThrAlaGlyAenProCysLeuGluGluAenAspValValLysPheTyrThr	549
Db	1024	ATCGCGCGCGCAGCAGGACTCTCAG---CAGCAACGATCCAAAGCACTGGACGTTGTA	1080	Db	1861	-----AATCCAGAGTTGGCAATGGAATATGATCAAT-----AAG	1896
Qy	193	MetThrGlnLysValArgTyrAlaProPheValAenGluLysValGlnPheAla	212	Qy	550	GluLeuIleGlyGlyCys-----LysPheArgGlyIleArg	561
Db	1081	CTG-----AGGAGTCACTTCTCGAATATGTCATC-----	1113	Db	1897	AAATGGTGTGGAGCTGTGTTCAAAGTGCAGCTTCGCTGAAATTTTTCACGCATGCAT	1956
Qy	213	LysAenPheValTyrAspAenAenSerIleLeuArgValProGluSerPheHisAspPro	232	Qy	562	IleGlyAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAenGlu	591
Db	1114	-----GTTTCTCGGTGCTTCTACTCCACT	1137	Db	1957	ATTGATGCC---GTGCACAGACTATCGCGCAACTAGTTTATACATGCAATGCTATTGGC	2013
Qy	233	AenArgPheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyr	252	Qy	582	TyrAlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaAla	601
Db	1138	ATGTTGGCCGCAAA---GACATTGGTGTGGCTGGAATGCTGGAAAGGATACTAT	1191	Db	2014	ATGGTTTTCATGAATGAATGCCAGATAGAGGTGGGGTCAAGTGTCTCCTAATTAACATCGAA	2073
Qy	253	IleGlyIleLysGluLeuPheAspGlyGluProValLeuAenPheAlaIleValAspLys	272	Qy	602	ThrGluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPhe	621
Db	1192	CAGAGCTCGGCCCACTCAGATGGACTCTCA---TTGAACATAGACATATCCTCTACC	1248	Db	2074	GCTGCCCTGAGCAACATTACACAGGGCTCCTCAA-----CTCCAGCTG	2118
Qy	273	LeuPheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeu---LeuIleValAsp	291	Qy	622	IleIleIleSerLysArgGlnLeuAenAla---TyrGlyPheValLysHisTyrCysAsp	640
Db	1249	CCATTCTTCAAA---CCTATCAGCGTGGTAGAATATGTCAAGAATTGCTGGGCACA	1302	Db	2119	CTCATTTGATTTCTCCAGATGTTAATGGGTATTATGAAGAATTAAGAGGGTGTGTGAG	2178
Qy	292	ProGlnSerCysAen---AspAspValArgLysAspLeuLysThrLysLeuMetAlaGly	310	Qy	641	HisThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAla	660
Db	1303	CCTACTAATGCTAATGGCCCTGACCTAGCGGCTCTTTCT-----	1344	Db	2179	ACTGAATTTGGATAGTATCCAGTGCCTCAAGCCA-----GGCCGAGCTCTTG	2229
Qy	311	LysMetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAenLys	330	Qy	661	SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAen	680
Db	1345	-----GACATTGATCGCTGAGAGTTAAGAAGCACTACGGGAGTTGCT	1389	Db	2230	AGCTTA-----GACAGGCAGTTCTCTGGAAAATGTCTCTCACTCAAAATCAAT	2274
Qy	331	LeuLysCysAlaGluValTrpAspAenGluMetSerArgLeuThrGluArgHisLeuThr	350	Qy	681	AlaLysLeuGlyGlyIleAenGlnGluLeuAspTrpSerGluIleAlaGluLysSerPro	700
Db	1390	GTTGAAACAACACACAGCGGAAGAGCAGCAAGTAGCAAGTCACTACG-----ATTACA	1443	Db	2275	GTCAGGCTGGAGGACGCAACTCAGTTCTT-----LeuThrMetTyrVal	715
Qy	351	PheLeuAepLeuCysGluGluAenSerLeuValTyrLysValThrGlyLysSerAspArg	370	Db	2305	-----CAGAGACCTTTGTACCCGGTGGCTTGAACACACAATAATTTT	2352
Db	1444	TCAGGCCACTGAGTCAGCTGAAC-----TTTTCTATGATGGAACTACCCAG---	1491	Qy	716	GlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaVal	735
Qy	371	GlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGluAenLys	390	Db	2353	GGTCCGATGTCAACCATCTCTTCTCGAGAGGACTCATCGGCTCGATCGCAGCTGTG	2412
Db	1492	-----ACTGTTATTCAGTACTTCTCGCAGCGGTACAAA	1524	Qy	736	ValAlaSerIleAen---ProGlyGlyThrIleTyrArgAenMetIle-----	750
Qy	391	Phe---IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyr	409	Db	2413	GTGGCTTCATGGAGCTGCCTGAGATCACCAAGTACAAAGCCCTCGTCTCTGCCAGCCA	2472
Db	1525	TACAGCTGCAGTACAGCTCTTGGCCCTGCTGCAATCCGCAACCTTCTAACCCTATA	1584	Qy	751	-----ValThrGlnGluCys	756
Qy	410	AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAenArgIle	429	Db	2473	CCTGGCAGGAGATTATACAAGATCTTTCACCATGACTGAAGTTGCCGAGAAATGCTGAT	2532
Db	1585	TATTTGCCAATGGAGTATGCACCAATTGTAGAA---GGCAAAAGATACTCCAAGAAGCTC	1641	Qy	757	ArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLys	776
Qy	430	AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys	449	Db	2533	GCTCCAGCACAGAAG---GCTGAAGTTTCAAGAGAAATTCATATGCGCGGCAATG	2586
Db	1642	ANTGACAAACAGTCACTGGCTCTCTGAGAGCAACATGCCAGCTCCCCAGAAAGGGAG	1701	Qy	777	PheValLysLeuLeuArgGluPheAlaGluAenAenAspAenArgAlaProAlaHisIle	796
Qy	450	GluAenThrLeuLysMetLeuLysGluLeuAepPheSerSerGluGluLeuAenPheVal	469	Db	2587	TTCAAGGAGTTGCTTATGCTCATTTCTACAGTAAGAAATGCTAAACGTAAGCTCAAGGATA	2646
Db	1702	CAGAAAATCATTTGAGATGTTTCAACACACAACACTACCCGCTGATAAGGTG-----GTG	1755	Qy	797	ValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeu	816
Qy	470	GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys	489	Db	2647	ATATTTTACAGGATGGTGAAGTATGAGTATGACAAATTCCTTCTGCTCTATGAGATG	2706
Db	1756	ACTGATTTTAGAATTAATTTTCCAAATCAGATGGCCACTATGCCAGCTGCGGTGCTGCT	1815	Qy	817	ArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPro---	835
Qy	490	GluProMetLeuValAenSerValAenGluGlnIleLysMetThrProValIleArgGly	509	Db	2707	GACGCAATCAAGAGGCTATTGCTCTTTG-----GACCCAGCA	2745


```
QY 416 uGluValHisGluLysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspIly 436
Db : : : : :
1517 CAANAATAGTGGAA---GGCAGAGATACCTCCAAGAGATTAAATCAGAATCAGATAAGAGC 1573

QY 436 sPheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAsnThrLeuLysMetIle 456
Db : : : : :
1574 TCTTTGGAGGACATGCCACCCACGTCATCGGAGCGTGATATTATTAAAGATGCT 1633

QY 456 uLysGluLeuAapPheSerSerGluLeuAenPheValGluArgPheGlyLeuCysSe 476
Db : : : : :
1634 TAAACATAACGCTTATCAGATGATCCT-----TACGCCAAAGAGATTGGCATTAAGAT 1687

QY 476 rLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeu---ValAs 495
Db : : : : :
1688 AAGTGATCTCTGGGCATCAGTAGAGGCACGAATTTTACCGGCTCCACGGCTTAAAGTACAA 1747

QY 495 nSerValAenGluGlnIleLysMetThrProValIle-----ArgGlyPheGlnG1 512
Db : : : : :
1748 TGAGACTGCTCGAGAGAGAGATTGCTTACCAAGAGTTGGTCAATGGAATATGATGAACAA 1807

QY 512 uLysGlnLeuAenValValProGluLysGluLeuLysCysAlaValPheValValAenG1 532
Db : : : : :
1808 GAAATATGGTAAATGGTGTAAAGTCAAGAGCTGGATGTCGTCGAATTTTGGCCGTAAT-- 1865

QY 532 uThrAlaGlyAenProCysLeuGluGluAenAspValValLysPheTyrThrGluLeu1 552
Db : : : : :
1866 -----GTGCGGAAAGTGTGTGTCGGGTTCTGCCATGAACCTGC 1906

QY 552 eGlyGlyCysLysPheArgGlyIle-----ArgIleG1 563
Db : : : : :
1907 TCTGATGTCTCAAGCTCAGGAATGGATTTTCGCTCCAGAACCTATCCTTCCACCATAA 1966

QY 563 yAlaAenGluAenArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAenGluTyrAl 583
Db : : : : :
1967 TGCACATCCAGATCAA-----GTGGAGCGTGTCTTAAAGAGCTAGGTAT-- 2009

QY 583 aPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaAlaThrG1 603
Db : : : : :
2010 -----CATGA 2014

QY 603 uAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIle1 623
Db : : : : :
2015 TCGAATGAACGTTCTTGA-----CCCCAGCGCAGGAACTTGATTGCTTATTGGGAT 2068

QY 623 eLleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHisThr1 643
Db : : : : :
2069 ACTACCTGATAACAATGGCTCGCTTTATGGTGAATTGAAGCGCGTGTGTGAATAGATCT 2128

QY 643 eGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAleSerLeuAr 663
Db : : : : :
2129 TGGAAATAGTTTCACATGCTGTGTGACGAGCAGGAGGTGTCAAAATG----- 2174

QY 663 gHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLysle 683
Db : : : : :
2175 -----AACAAACAATCTTCGCAATCTGCTCTGGAAGATAAATGTCAAGT 2221

QY 683 uGlyIleAenGlnLeuAenAspTrpSerGluIleAlaGluIleSerProGluGluLyl 703
Db : : : : :
2222 TGGGGGCGAGAACACTGTGCTG-----GTGGATGCTGTGTCA----- 2258

QY 703 sGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyLeAs 718
Db : : : : :
2259 -----AGCGTATTCCTCTGGTAACCGACAGACCTCAATATATTATTTGGTGCTGA 2308

QY 718 pValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValValAlase 738
Db : : : : :
2309 TGTACCACATCTCATCTCGGAGAGGACAGCAGTCCCTCAATTGCTGTGTTGTAGCCTC 2368

QY 738 rIleAen---ProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluGluCysAr 757
Db : : : : :
2369 CCAAGATTGGCCTGAGGTGACAAAGTATGCTCGGTTGGTTTCTGCTCA----- 2417
```

```
QY 757 gProGlyGluArgAlaValAlaHisGlyArgGlu----- 768
Db : : : : :
2418 -----GCCACCGACAGAGCTGATAGAGATCTATATAAAATCTCG 2458

QY 769 -----ArgThrAspIleGluAlaLysPheValLysLeuLeuArgGluPh 784
Db : : : : :
2459 GCAGGATCCACAGAGAGGAACAGTTAGTGGTGGCATGATCCCGTGGAGCTCTTATATCCTT 2518

QY 784 eAlaGluAenAenAspAenArgAlaProAlaHisIleValValTyrArgAspGlyIle 804
Db : : : : :
2519 CAAAAGATCAACTGCTGAGAAG---CCCCAGCAATAATATTTTACAGGATGGCGCTTAG 2575

QY 804 rAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluVally 824
Db : : : : :
2576 TGAAGCCCAATTTTACCAAGTTCTACTTATGAATTAATGAATCAATCCGAAAGCATGTGC 2635

QY 824 eGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleVal1 844
Db : : : : :
2636 CTCCTCTGGAGACA-----AATTACCAACCAAGGTGACTTTTCATTGTGCT 2680

QY 844 eGlnLysArgHisAenThrArgLeuLeuArgMetGluLysAspLysProValValAs 864
Db : : : : :
2681 TCAGAAACGTCACCAACACAGATTTATTCACATAATCAACAGCATCAGAACTCAGTTGA 2740

QY 864 nLysAspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGlu1 884
Db : : : : :
2741 CAGG----- 2744

QY 884 uAspMetLysGluSerLysGluThrGlyIleValAenProSerSerGlyThrValAs 904
Db : : : : :
2745 -----AGCGGGAACATACTCCCT-----GGTACGTTGTAGA 2776

QY 904 pLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHisHisGlyValle 924
Db : : : : :
2777 TTCAAGATCGCTCATCCAACTTGACTTGACTTACCTGTGTAGCCATGCTGGCATTA 2836

QY 924 uGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAs 944
Db : : : : :
2837 GGGTACTAGTCTCCAGCTCATTTATCATGCTCTGTGGGATGAACAACTTCACAGCTGA 2896

QY 944 pGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysPro1 964
Db : : : : :
2897 TGCATTGCAGATCTTACCAACAACTTTGCTACACCTATGCAAGGTGCACTCGCTCTGT 2956

QY 964 eSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTy 984
Db : : : : :
2957 ATCAATTGTTCCACCTGCTTATATGCTCATCTGCTGCCCTCCGTCCT----- 3005

QY 984 rArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGln-----ProArgTh 1000
Db : : : : :
3006 -CGTTTCTATATGGAACCCAGATACATCTGACAGCAGCTCTGCTTAGTGGGCTGGTCT 3064

QY 1000 rArgHisGluMetGluHisPheLeuGlnThrAenValLysTyrProGly 1016
Db : : : : :
3065 ACGTGGGCACTTTCTGCTCATCAACATCACGTACTCGGCCCCCTGGT 3113
```

RESULT 8

```
US-11-218-305-22486
; Sequence 22486, Application US/1121805
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Mc Laird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
```



```
Db 2518 -----AACAGCAGCAGTATCTTGCAATGTTGCCTGAAATAATATGTGAAG 2565
QY LeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGluIleSerProGluGlu 702
Db 2566 GTTGGGGAAGGAATACTGTGCTT-----ThrMetTyrValGlyIle 717
QY 703 LysGluArgArgLysThrMetProLeu-----ThrMetTyrValGlyIle 717
Db 2599 TTGGCAAGGAGA-----ATCCCCCTTGTCAGTGACATAGCAGCTATATCTTTGGTGT 2652
QY 718 AspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAla 737
Db 2653 GATGTGACCCATCCCATCTCGGGAAGATTCTAGTCCTTCCATTCGACGTGGTGTGT 2712
QY 738 SerIleAsn---ProGlyGlyThrIleTyrArgAsnMetIleValThrGln----- 753
Db 2713 TCTCAAGACTGGCCTGAGGTTACAAAGTATGCAAGGATTGTTGAGTGCTCAAGCCCATCGT 2772
QY 754 GluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeu 773
Db 2773 CAAGAAATTCATACAGATCTTTCAAGGTATGCAAGATCCCGAAGGGGAGCTGTCTCT 2832
QY 774 GluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaPro 793
Db 2833 GGTGGCATGATCAGGAGCTTCTCATATCTTTCTGGAGGGCAACTGGACAGAAA---CCA 2889
QY 794 AlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHis 813
Db 2890 AAGAGGATCATATTTACAGAGGATGGCGTCAGTGAAGGACAATTTACCAAGTTCTGTGTG 2949
QY 814 AspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGlu 833
Db 2950 TATGAACCTTGATGCCATTAGAAGGCTGTGGCTCATTTGGAGTCT----- 2994
QY 834 AspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeu 853
Db 2995 GACTACCAGCCTCCAGTTACTTTTGCCTGGTCCAGAGCGTCATCACACAGGTTGTTT 3054
QY 854 ArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAsp 873
Db 3055 GCTAATAATACAAATGATATCGTGTGTCGATAAA----- 3090
QY 874 ValAlaValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGly 893
Db 3091 -----AGCGGG 3096
QY 894 IleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPhe 913
Db 3097 AACATACTGCCT-----GGCACCGGTGGTGGACTCGAAGATCTGCCATCCAATGAGTTT 3150
QY 914 AspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThr 933
Db 3151 GATTTCTACCTGTGCAGCCATCTGTCATTTCAGGGAACAAGCGCCCTCCCATACCAT 3210
QY 934 ValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeu 953
Db 3211 GTTCTGTGGATGAGAACACTTTACCGCTGTGTTGGTTCGAACCTCTCAACAACTGTG 3270
QY 954 AlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAla 973
Db 3271 TGTATACCATGATGCTAGGTGCACACGCTCAGTATCGATTGTTCTCTCATACATATGCT 3330
QY 974 HisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGly 993
Db 3331 CACCTGGCAGCCTTCGAGCT-----CGGTTCTCATGTGAGCCAGATACGAGT 3378
QY 994 AspTyrAlaGlnProArgThrArg 1001
Db 3379 GACAGTGGATCTATGCAAGCCGT 3402
```

RESULT 9

US-11-266-748A-25558

```
; Sequence 25558, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 25558
; LENGTH: 1793
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-25558

Alignment Scores:
Pred. No.: 1,736-40 Length: 1793
Score: 493.00 Matches: 160
Percent Similarity: 43.7% Conservative: 89
Best Local Similarity: 28.1% Mismatches: 202
Query Match: 9.2% Indels: 120
DB: 8 Gaps: 18

US-10-645-746-3 (1-1020) x US-11-266-748A-25558 (1-1793)

QY 434 GlnAspLysPheLeuLysArgAlaThr---ArgLysProHisAspTyrLysGluAsnThr 452
Db 3 CAGACCTCAACCATGATCATGAGCAACTCTAGTCCGCCCGCCGATCGGCAAGAAGAGATT 62
QY 453 LeuLysMetLeuLysGluLeuAspPheSerSerGluLeuAsnPheValGluArgPhe 472
Db 63 AGCAATTTGATGTCGAAGTGCAAGTTTCAACACAGATCCA-----TACGTCCGTGAATTT 116
QY 473 GlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMet 492
Db 117 GGAATCATGTCGTCGAGATGAGATGACAGACGTGACTGGCGGGGTGCTGCAGCCGCCCTCC 176
QY 493 LeuValAsnSerValAsnGluGlnIleLysMetThrProVal-----IleArgGlyPhe 510
Db 177 ATCTCTTACGGGGGCGAGGAATAAGCTATTTCGACCCCTCTCCAGGGCGTCTGGGACATG 236
QY 511 GlnGluLysGlnLeuAsnVal---ValProGluLysGluLeuCysCysAlaValPheVal 529
Db 237 CGGAACACAGCAGTTCCACACGGGCATCGAGATCAAGGTGTGGGCCCATTCGTGCTTCGCC 296
QY 530 ValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAspValLysPheTyrThr 549
Db 297 CCCCAG-----CGCCAGTGCACGAAGTCCAT-----CTGAAGTCTTCCACA 338
QY 550 GluLeuIle-----GlyGly-CysLysPheArgGlyIleArgIleG1 563
Db 339 GAGCAGCTCAGAAAGATCTCGAGAGACGCTGTCATGCCCATCCAGGCCCGCGCTGCTTC 398
```

```
QY 563 YAlaAenGluAenArgGlyAlaGlnSerIle-----MetTyrAspAlaThrLysAenGI 581
Db 399 TGCAAAATACCGC-CAGGGGGCGGACGCGTGGAGCCCATGTTCGGCACCTGGAAGAACAC 457
QY 581 uTyrAlaPheTyrLysAenCysThrLeuAenThrGlyIleGlyArgPheGluIleAlaA 601
Db 458 GTATGCG----- 464
QY 601 aThrGluAlaLysAenMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPh 621
Db 465 -----GGCCTGCAGCTGGT 478
QY 621 elleIleSerLysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHi 641
Db 479 GGTGTCATCTCTCCCGGCAAGACGCGCGTGTACGCCGAGGTCAAGCGCGTGGAGACAC 538
QY 641 sThrIleGlyValAlaAenGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSe 661
Db 539 GGTGCTGGGATGGCCACGCGAGTCGCGTCAGATCAAGAAGCTG----- 581
QY 661 rLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAenAl 681
Db 582 -----CAGAGGACCAACGCGCACAGACCTGTCCAACCTCTGCGCTGAAGATCAACGT 631
QY 681 aLysLeuGlyGlyIleAenGlnGluLeuAspTyrSerGluIleAlaGluIleSerProGI 701
Db 632 CAAGCTGGGAGGCGTGAACAACATCTCTG-----CTGCCCA 667
QY 701 uGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyIleAspValThrHi 721
Db 668 GGCAGCGCGCGGTGTTCAGCAGCCGCTC-----ATCTTCTGGAGCAGACGTCACTCA 724
QY 721 sProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValAlaSerIleAenPr 741
Db 725 CCCCCCGCGCGGATGGGAAGAAGCCCTCCATTGCGCGCGTGTGGCAGCATGAGACG 784
QY 741 oGlyGlyThrIleTyrArgAenMetIleValThrGlnGluGluCysArgProGlyGluAr 761
Db 785 CCACCCCAATCGCTACTGCGCCACCGTGGCGGTGCAGCAG----- 824
QY 761 gAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLe 781
Db 825 -----CACCGCAGAGATCATACAGACCTTGGCGCCCATGTGTCGCGAGTCTCT 874
QY 781 uArgGluPheAlaGluAenAenAspAenArgAlaProAlaHisIleValValTyrArgAs 801
Db 875 CATCCAGTCTACAGAGTCCACGCGCTCAAG---CCACCCGCGATCATCTTCTACCGCGA 931
QY 801 pGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSe 821
Db 932 CGGTGTCTCTGAAGGCGAGTTCACGAGGTCTCCACACGAGTGTGTCGCCATCCGTGA 991
QY 821 rGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProLysTyrThrPh 841
Db 992 GGCCTGTATCACTA-----GAAAGAAGACTACACCGCCGGGATCACTT 1036
QY 841 elleValIleGlnLysArgHisAenThrArgLeuLeuArgArgMetGluLysAspLysPr 861
Db 1037 CATCGTGGTGACAGAGGACCAACCCCGGCTCTTC---TGCACTGACAGAAGACGCG 1093
QY 861 oValValAenLysAspLeuThrProAlaGluThrAspValAlaValAlaValLysGI 881
Db 1094 GGTGGGAAAGTGGAAACATCCAGCA----- 1121
QY 881 nTrpGluGluAspMetLysGluSerLysGluThrGlyIleValAenProSerSerGlyTh 901
Db 1122 -----GGCAC 1126
QY 901 rThrValAspLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHisHi 921
Db 1127 GACTGTGGACACGAAATATCACCCACCCCGAGTTCGAGTCTTACCTCTGTGTAGTCAACG 1186
QY 921 sGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMe 941
```

```
Db 1187 TGGCATCCAGGGGACAGCAGCGCTTCGCACATATCACGTCTCTGGGACGACAAATCGTTT 1246
QY 941 tSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaAArgCysAr 961
Db 1247 CTCTCTGTGATGAGCTGCAGATCCCTAACCTACCACTGTGTGCACACCTACGTGCGCTGCAC 1306
QY 961 gLysProIleSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLy 981
Db 1307 ACGTCTCGTCTCCATCCAGCGCAGCATACTAGCTCACCTGTGTGGCTTCCTCCGGCCAG 1366
QY 981 sGluLeuTyrArgThrTyrLysGluHis 990
Db 1367 G---TACCACCTGGTGGATAAGAACAT 1391
```

RESULT 10

```
US-10-449-902-19320
; Sequence 19320, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19320
; LENGTH: 2503
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK069685
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-19320
```

Alignment Scores:

Pred. No.:	1,46e-39	Length:	2503
Score:	486.00	Matches:	199
Percent Similarity:	40.6%	Conservative:	116
Best Local Similarity:	25.7%	Mismatches:	289
Query Match:	9.1%	Indels:	171
DB:	6	Gaps:	32

US-10-645-746-3 (1-1020) x US-10-449-902-19320 (1-2503)

```
QY 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyIleLysGlu 257
Db 220 TCGAGGGAATCTGCACAAAGGCGCATCTTTGCGTGAAGGGGTTCTACAGAGTTCGCCGGTC 279
QY 258 LeuPheAspGlyGluProValLeuAenPheAlaIleValAspLysLeuPheTyrAenAla 277
Db 280 ACGCAGCAGCGCTTCTCT---CTGAACATAGACATGTCTTCGACTGCTTTC----- 327
QY 278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 297
Db 328 -----ATTGACCTGCTGGGTGCTGAAT 351
QY 298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
Db 352 TTGTAGAAAAGACTATTGGACGTCGAATC---ACTAATGCTATTACTGTGGGATATTTT 408
QY 318 AlaArgProArgIleArgGlnLeuLeuGluAenLeuLysLysCysAlaGluValTyr 337
Db 409 TTGAACAATATTGGGAATGAATTGATGAGACCCCTTAAGGCTGTTAAGTTGAATC--- 465
```

338 AspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGlu 357
 466 ---ACTCACCAGGAAATCTACCGAAGAAGTAC---AGCATTTGCTGGCTTCACCCGACG 519
 358 AsnSerLeuValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyr 377
 520 TCTCGAGATGTTCAAGCGTTCCACA---TCATCTCATGGTATCAAGACTGTCAAGAGGATAT 576
 378 AspThrThrLeuPheLysIleTyrGluGluAsnLysLysPhe-----IleGluPhe 394
 577 -----TTCAACAACAAAANTACATCTGAAGTTAGCTTTT 609
 395 ProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyrAlaValPrometGlu 414
 610 GGTATCTTCATCGCTGCAAGTTGGCAGCAAGGAGAGCCGAATTAACCTGCCCATGGAG 669
 415 HisLeuGluValHisGluLysProGluArgTyrLysAsnArgIleAspLeuValMetGln 434
 670 CTTTGC AATATA---GTTCTCGGACACAGATACAAAGAACAGGCTCACTCCGACACAGGCT 726
 435 AspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAsnThrLeuLys 454
 727 TCCAACTCATTAACATAACCAACGATCGTCTTGTGCAC---CGTAGAGCTCCATTCGT 783
 455 MetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheValGluArg----- 471
 784 CAG-----ACTGTTAGCAGCAACCCAGTATAACAGTACGGAAACGCGCAGATGAG 831
 472 PheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluPro 491
 832 TTTGGCATAGAAAGTTGATCTCTTATCTCTACTTTTAAAGGCTAGAGTTTGAAGACTCCA 891
 492 MetLeuValAsnSerValAsnGluGlnIleLysMet---ThrProValIleArgGlyPhe 510
 892 ATGCTGAAGTACCATGATTTCTGGAAGGGTGAGAGTATGCACGCCACAGGATGGGGCGTGG 951
 511 GlnGluLysGlnLeuAsnValVal-----ProGluLysGluLeuCysCysAlaVal 527
 952 AACATGAAGAGCAAGAAAGTAGTTAAACGGTGCTACAATTAAAGCTGGCGATGT----- 1005
 528 PheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAspValValLysPhe 547
 1006 -----GTCAACTTGTGCGAGGGT-----TTGGATAATCGTGTGTGAAGCAATTC 1050
 548 TyrThrGluLeuIleGlyGlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsn 567
 1051 TGCCTTCAATTGGTCAGAACGTCCAAATAACTGCGACTGGAC----- 1092
 568 ArgGlyAlaGlnSerIleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsn 587
 1093 -----TTTGC GAAT 1101
 588 CysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMet 607
 1102 GTGAGCCTT-----CCAAATATTGAAGCTGATCTCTCAATATGTT 1140
 608 PheGluArgLeuPro-----AspLysGluGlnLysVal 618
 1141 AAAACTGATCTTCTATGCGGTATCAGAGACGATGCAGCTGGTCGAGGGATAACCAAGATT 1200
 619 LeuMetPheIleIleIle-----SerLysArgGlnLeuAsnAlaTyrGlyPheVal 635
 1201 GACCTCTCTACTTGTGTAAATACACAGATCATATAAATAATGCCAGCTTATATGGTGACGTT 1260
 636 LysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVal 655
 1261 AAAAGAATCTGTGAACACAGAAATCGGTGTATTGTACACAGTGTGTTCAGCGCAAGCAAGTC 1320
 656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyr----- 673
 1321 TACAAG-----GAGAGGAATGTTCTAGTACTGCGCA 1350
 674 GlnIleAlaLeuLysIleAsnAlaLysLeuGlyGly-----IleAsnGln 688

RESIT.T 11

US-11-218-305-543
: Sequence 543. Application US/11218305

Publication No. US20060141495A1

REGISTRATION NO. 00200014112581
GENERAL INFORMATION:
APPLICANT: MONSANTO TECHNOLOGY, LLC

APPLICANT: McLaird, Paul

APPLICANT: Tao, Nengbing

APPLICANT: Wu, Kunsheng


```

, TITLE OF INVENTION: Corn.
, FILE REFERENCE: 38-21 (53660)B
, CURRENT APPLICATION NUMBER: US/11/218,305
, CURRENT FILING DATE: 2005-09-01
, PRIOR APPLICATION NUMBER: US 60/606,880
, PRIOR FILING DATE: 2004-09-01
, NUMBER OF SEQ ID NOS: 25043
, SOFTWARE: PatentIn version 3.2
, SEQ ID NO 543
, LENGTH: 5408
, TYPE: DNA
, ORGANISM: Zea mays
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (766)..(767)
, OTHER INFORMATION: n is a, c, g, or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (769)..(769)
, OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-543

```

Alignment Scores:	
Pred. No.:	2.32e-38
Score:	479.00
Length:	5408
Matches:	153
Percent Similarity:	44.3%
Best Local Similarity:	27.3%
Conservative:	95
Mismatches:	196
Query Match:	9.0%
Indels:	116
DB:	7
Gaps:	22

US-10-645-746-3 (1-1020) x US-11-218-305-543 (1-5408)

	QY	470	GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys	489
	DB	2082	CAGCAATTTGGGATTTCTTGGGTACGAATCACAGACAGCATCGGTCAGGCATCTCCCT	2141
	QY	490	GluProMetLeu---ValAenSerValasnGluGlnIleLysMetThrProValIleArg	508
	DB	2142	CCACCAAACCTCAAACCTTGCGGCATCCAAATGGGCAGACCTCCAATTGAGTATCGATCAT	2201
	QY	509	GlyPheGln-----GluLysGlnLeuAsnValValProGluLysGluLeuCys	524
	DB	2202	GGCTGCCAGTGGAAATCTTGTGAAGAGAGACTA----GTAGAGGCCCGGTCTTCAG	2255
	QY	525	Cys-----AlaValPheValValasnGluThrAlaGlyAsnProCysLeuGluGluAsn	542
	DB	2256	TGCTGGGCATCTCGCACTTCAGTGTCTGAGCCGCTCTGGCTCTGGCGCCGTCAGGAGCCC	2315
	QY	543	---AspValValLysPheTyThrGluLeuIleGlyGlyCysLysPheArgGlyLleArg	561
	DB	2316	CTCGATACAAGGATGTTTGTGTGAGAAGATTGTGAGGAAGTGCTGTGAGCTTGGTATCCGT	2375
	QY	562	IleGlyAlaAsnGlu-----AsnArgGlyAlaGlnSerIleMetTyRaspAlaThr	578
	DB	2376	ATGAACCCCTAATCCATGCTTTCGTGCACATAACAAGGATGGCAGTGCTCTTCGATCCACAT	2435
	QY	579	LysAsnGluTyRalaPheTyLysAsnCysThrIleuAsnThrGlyLleGlyArgPheGlu	598
	DB	2436	-----GGACTATCAAGACGCTAAACAAGCAAAA	2465
	QY	599	IleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysVal	618
	DB	2466	CAAGCTGCAGTG-----AGCAAGAAGCAGAGGTTG	2495
	QY	619	LeuMetPheIleIleIleSerLysArgInLeuAsnAlaTyrglyPheValLysHisTyr	638
	DB	2496	CAGCTCCTTTCTGCCCCGATGTCGGCAGCAGCATTCGGGGTACAAAGACACTGAAGCTGATT	2555
	QY	639	CysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAla	658
	DB	2556	TGTGACACACAGCTGGGGATCTTCACCGCTGTTTCTGTAGCCGACCGCGCA-----	2606
	QY	659	LeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrglnIleAlaLeuLys	678

2607	Db	-----AACATCGAAGGGGACAGGACGATACATGACGAATCTTGCTTAAAG	2654
679	Qy	IleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaGluIle	698
2655	Db	ATCAACGGCAAGCTTTGGGGCAGCAAGTTCCAGTG---TTTGACTCGCTCCCAACGGGTC	2711
699	Qy	SerProGluGluLysGluAArgLysThrMetProLeuThrMetTyrValGlyIleAsp	718
2712	Db	GGT-----GGTGGGTACCTTTC---ATGTTTCATCGGTGCTGCAC	2747
719	Qy	ValThrHisProThrSerTyrSerGlyIleAspTyr-SerIleAlaValValAlaSer	738
2748	Db	GTTAACACACCGTCCCGGTAACTGGAGAGCCCATCGATCGAGCCGTGGTGCCTCT	2807
739	Qy	IleAsnProGlyGlyThrIleTyrAArgAsnMetIleValThrGlnGluCysAArgPro	758
2808	Db	GTCAACTCTGGT-----GTCAACAAGTATGTGACCAAGATCCGTGCCACCGCG	2855
759	Qy	GlyGluArgAlaValAlaHisGlyAArgGluAArgThrAspIleLeuGluAlaLysPheVal	778
2856	Db	-----CACCGCTGGAGGTGATTCAG---CAGCTTGGT	2885
779	Qy	LysLeuLeuAArgGluPheAla-----GluAsnAsnAspAsnAArgAlaProIleHis	795
2886	Db	GAGATCTCGCGGAGCTCATTTGGAGTCTTTGAGAACAGACCGCTGAACGCGCAGAG	2945
796	Qy	IleValValTyrAArgAspGlyValSerAspSerGluMetLeuAArgValSerHisAspGlu	815
2946	Db	ATCATCTACTCTCGTATGGCGTGGAGCGAGCGGCGAGTTCCGATATGCTCTGAAAGAGGAG	3005
816	Qy	LeuAArgSerLeuLysSerGluValLysGlnPheMetSerGluAArgAspGlyGluAArgPro	835
3006	Db	CTGGCTGACCTGGAGAGCGCATCAAGTG-----AATGGCTAT	3044
836	Qy	GluProLysTyrThrPheIleValIleGlnLysAArgHisAsnThrAArgLeuAArgAArg	855
3045	Db	GCGCCAACCATCACCTGGTGGTGGCGCAAGAGCGGACCAACACTCGCTGTTTCCCGCAGG	3104
856	Qy	MetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAla	875
3104	Db	-----	3104
876	Qy	ValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleVal	895
3105	Db	-----GACGACACGACGCGCGGTGGACCCGCTCGCTACGACTTC	3140
896	Qy	AsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhe	915
3141	Db	CCTCTCT-----GGCACGGTGTGGACACGGCGGTGGACCCGCTCGCTACGACTTC	3194
916	Qy	PheLeuAlaSerHisGlyValLeuGlyThrSerAArgProGlyHisTyrThrValMet	935
3195	Db	TACCTGTGACGCCACACTGGGATCTTGGGGAGAGAGCCCGCTACTACACCTTC	3254
936	Qy	TyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPhe	955
3255	Db	GTGGACGACGACGGCTTCGGCTCCGACGACCTGCAGAGCTGATCTACAACTGTGCTTC	3314
956	Qy	LeuSerAlaAArgCysAArgLysProIleSerLeuProValProValHisTyrAlaHisLeu	975
3315	Db	GTGTTTCGCGCGGTGCACCAAGCGGTGTGCTGTGGGACGCGCGCTCTACTACGCCACCTTC	3374
976	Qy	SerCysGluLysAlaLysGluLeuTyrAArgThrTyrLysGluHisTyrIleGlyAspTyr	995
3375	Db	CGCGCCCTACGTGGCAGG---CTCTAC-----TACGAGCGTCCCATGATGGCGTCCAG	3425
996	Qy	AlaGln-----ProAArgThrAArgHisGluMetGlu	1005
3426	Db	GCACGCGAGGGGGTCTTTTGACGTCCTAACTTCCCGAGCGCTGCACAAAGATGTGAG	3485

RESULT 12

US-11-218-305-19388

```

; Sequence 19388, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19388
; LENGTH: 3350
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (738)..(742)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908)..(930)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3235)..(3236)
; OTHER INFORMATION: n is a, c, g, or t
; US-11-218-305-19388

Alignment Scores:
Pred. No.: 1,08e-37 Length: 3350
Score: 469.50 Matches: 225
Percent Similarity: 38.4% Conservative: 171
Best Local Similarity: 21.8% Mismatches: 245
Query Match: 8.8% Indels: 245
DB: 7 Gaps: 41

US-10-645-746-3 (1-1020) x US-11-218-305-19388 (1-3350)

QY 20 GluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysLys 39
DB 283 GAACAAAGATGCCATTGAACCTAGGAGCTCTGGTCCAGAGAAATGGATTTCGCAGAAAG 342
QY 40 -----ValLeuLeuValAenTrpPheLysPheSerSerLysLysLysTyrAspArg 56
DB 343 GGGCAGCCAAATGAAGCTGATAACAATCACTTCAAAAGTTCTCTTGTGTAATGCTGAGAA 402
QY 57 GluTyrTyrGluTyrGluValLysMetThrLysGluValLeuAenArgLysProGlyLys 76
DB 403 TTTTCTACCATCTACTATGTCAATTGAAGTATGAA----- 438
QY 77 ProPheProLysLysThrGluLeuProLeuProAspArg-----AlaLysLeu 92
DB 439 -----GATGATACACCGGTT---GATCGCAAAAGGGTTCAGAAAGGAAGTG 480
QY 93 PheTrpGlnHisLeuArgHisGluLysLysGlnThrAspPheLeuLeuGlu----- 109
DB 481 ATT-----GAAAACTGCGACAAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522
QY 110 ----AspTyrValPheAspGluLysAspThrValTyrSerValCysArgLeuAenThrVal 128
DB 523 AAAGATTTTTCCTATGATGGTGCAGAGAGCCGTGTTTCAAAATGGTGTCTTCTCTCAAGTT 582
QY 129 ThrSerLysMetLeuValSerGluLysValValLys----- 140
DB 583 AAAATTGAGTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
QY 141 -----LysAspSerGluLysLysAspGluLysAspGluLysLysLysLysLysLysLys 157

```

```

DB 643 AGTCAGGCAATGACAGTCTCCCGAAGTGCAGGAAAGAGGTCGAGGCGCTTACAAT 702
QY 158 ThrMetIleLeuThrTyrArgLysLysPheHisLeuAenPheSerArgGluAenPro---- 176
DB 703 ACAAG-----ACCTAT-----AAGGTCGAGCTCTCTTTTGGCGCANNNTTCCCTATG 750
QY 177 -----GluLysAspGluGluAlaAenArgSerTyrLys 187
DB 751 AGTCAATCTCACAGCTTTGAGGGTCAGGAGTCAGACGACACTCAGGAAGCAATTCGA 810
QY 188 PheLeuLysAenValMetThrGlnLys-----ValArgTyr 199
DB 811 GTGATTGACATTATTCTGAGGAGCAGCTCAGCTAAGCAGGGTTGCCATTAGTAGAGCAA 870
QY 200 AlaProPheValAenGluGluLysValGlnPheAlaLysAenPheValTyrAspAen 219
DB 871 TCATTCTTCCACAACAAT-----CCTTCCAATTTTCTT---GACNN 909
QY 220 AenSerIleLeuArgValProGluSerPheHisAspProAenArgPheGluGlnSer--- 238
DB 910 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 969
QY 239 LeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyLysGluLeu 258
DB 970 CTTTCACTCAAGCCCGCGGTCGCTACAATGATAGTGAACCTGGT----- 1017
QY 259 PheAspGlyGluProValLeuAenPheAlaIleValAspLysLeuPheTyrAenAlaPro 278
DB 1018 -----CCTGTCTGACTTCTGCTTGCCAAATCAG-----AAAGTTGGT 1056
QY 279 LysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAenAspAep 298
DB 1057 GATCCAAGCATGATTGATTGG----- 1077
QY 299 ValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAla 318
DB 1077 ----- 1077
QY 319 ArgProArgIleArgGlnLeuLeuGluAenLeuLysLysCysAlaGluVal---Trp 337
DB 1078 ---GCTAAGGCCCAAGCGTGCTGAAGAACTTTGAGGATAAAATAAAGTCCAGCAACCAA 1134
QY 338 AspAenGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGlu 357
DB 1135 GAACAGAGATTGTTGGTCTCAGCAAGA-----ACTGTCTGTGAG 1176
QY 358 AenSerLeuValTyrLysValThrGlyLysSerAspArgGlyArgAenAlaLysLysTyr 377
DB 1177 CAATTATTTC-----ACACTGAAACATAAAATGTTAAACAATGGTGAATGCTGCTGAA 1224
QY 378 AspThrThrLeuPheLysIleTyrGluGluAenLysLysPheIle-----GluPhePro 395
DB 1225 GAGATCACTGTTTATGATTACTTCGTAAGCAGCGTGGCATAGTGTGCAATCTCTGGT 1284
QY 396 HisLeuProLeuValLysSerGlyAlaLysGluTyrAlaValProMetGluHis 415
DB 1285 GATCTTCTTGTGATCAATGTGGGAAACTAAGCGGCCCAACATATTTTCCAATTGAGTTA 1344
QY 416 LeuGluValHisGluLysProGlnArgTyrLysAenArgIleAspLeuValMetGluAen 435
DB 1345 TGCAGTCTT---GTCCCTTTACAAAGATACACTAAAGCTTTTGAACACGCTTTCAGAGGTC 1401
QY 436 LysPheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAenThrLeuLysMet 455
DB 1402 TCATCTCGTGGGAAATCTAGGCAGAAACCGCAGGAAAGGATGCTGTTTTATCTGATGTG 1461
QY 456 LeuLysGluLeuAspPheSerSerGluGluLeuAenPheValGluArgPheGlyLeuCys 475
DB 1462 CTGCAAGAGCACTATGATGACAGCCCATG-----TTGAGGCATGCGGATTACA 1515
QY 476 SerLysLeuGlnMetIleGluCysProGlyLysValLeuLysGluProMetLeuValAen 495

```

1516 ATTGCTAGAAATTTACAGAGTTGATGGTAGGGTATTGACGCCACCTTAAGCTTAAAGCT 1575
Db
496 SerValAsnGluGlnIleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeu 515
QY
1576 GGGAAATCGTGAAGACATTTTACACGCAATGGTAGTGAACCTTCAACAATAAGAGGCTC 1635
Db
516 AsnValValProGluLysGluLeuCyCyAlaValPheValValAsnGluThrAlaGly 535
QY
1636 -----ATTAGAGCTTGAGTGTCT----- 1653
Db
536 AsnProCysLeuGluAsnAspValValLysPheTyrThrGluLeuIleGlyGlyCys 555
QY
1654 -----GAGAAATCGCGGTGTAAACTTTTCTGACGA-----TGC 1689
Db
556 LysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyr 575
QY
1690 AATGTCAGGATCTTGTCCGGATCTCATCAAGTGTGGAGGATGAAGGCAATTATGTT 1749
Db
576 AspAlaThrLysAsnGluTyrAla---PheTyrLysAsnCysThrLeuAsnThrGlyIle 594
QY
1750 GATGCT-----CCTTTTGCTGTAATTTGATGAGAAATCCTCAATG----- 1788
Db
595 GlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLys 614
QY
1789 ---AGACGGTCACTGCTATAAGAGGGTTGAAGACATGTTTGAACAAGTGAACCTAAG 1845
Db
615 -----GluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAla 631
QY
1846 CTTCTCGGACCAAAAGTTTCTTTTGTGTCTTAGCTGAAGAGAAATTCGTGATTT 1905
Db
632 TyrGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThr 651
QY
1906 TATGGCCCTTGAAGAAGAAATCCCTGTGGAATTTGGGATCGTTACACAATGTGGCA 1965
Db
652 SerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIle 671
QY
1966 CCAACTAGAGTG----- 1989
Db
672 PheTyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnGluLeuAsp 691
QY
1990 CTTACAAATGTCTACTTAAGATAAAATGCAAGCTGGTGGCATGAATTCGTTGCTCAA 2049
Db
692 TrpSerGluIleAlaGluIleSerProGluLysGluArgLysThrMetProLeu 711
QY
2050 -----ATTGAACATCCCCA-----GCAATTCCTCTT 2076
Db
712 -----ThrMetTyrValGlyIleAspValThrHisProThrSerTyrSer 726
QY
2077 GTATCCAAGTCCCAACTATAATCTTGGGAATGGATGTGTACACGGTCTCTCGACAT 2136
Db
727 GlyIleAspTyrSerIleAlaValAlaSerIleAsn---ProGlyGlyThrIle 745
QY
2137 TCTGATGTACCATCTATTGCTGTGTGTGTGTCTCGTGAATGGGCTCTTATCTCGAA 2196
Db
746 TyrArgAsnMetIleValThrGlnGluCyCyArgProGlyGluArgAlaValAlaHis 765
QY
2197 TACAGAGCTTCTGTCCGACCCCAATCCTCAAAATGGAATGATTCGACTCATTTGTTAAG 2256
Db
766 GlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAla 785
QY
2257 CCACGGGAAGCTCAAGAT-----GATGCTGTATCCGGAGTGTCTGATTGACTTCTAC 2310
Db
786 GluAsnAsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAsp 805
QY
2311 ACCAGTTCTCGGAAGAAGACCTGACCAAGTCAATCATATTACGGACGGTGTAGCGAA 2370
Db
806 SerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGln 825
QY
2371 AGTCAGTTTAATCAGGTGTGAACATTTGATGTGAACAATC---ATCAGAGCTTGCAAA 2427
Db
826 PheMetSerGluArgAspGlyGluAspProGluProGluProGluPheIleValIleGln 845
QY
2428 TTTCTTGATGAGAA-----TGAATCCCAAGTTACGTTGATTAATGCCCCAG 2475
Db

846 LysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLys 865
QY
2476 AAGAATCATCACACT-----AAA 2493
Db
866 AspLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAsp 885
QY
2494 TTTTTCATTCCTCGAAGCCAGAT----- 2517
Db
886 MetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLys 905
QY
2518 -----AATGCCACAGGAATCTGTGTGGACAAAC 2547
Db
906 LeuIleValSerLysTyrLysPheAspPheLeuAlaSerHisHisGlyValLeuGly 925
QY
2548 AAAGTCTGCCATCCAAAGAACTTCGATTTCTACATGTGTGGCATCTGTGGAATCATCGG 2607
Db
926 ThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGlu 945
QY
2608 ACTACAGGGCCAACTCACTACCACTCTCTCATGATGAGATAGGCTTTCAGTCTCTGATGAT 2667
Db
946 ValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSer 965
QY
2668 CTCAGAGCTGTGTGATTCGCTCTCTATGTGTACCAAGGACCAACAGCCATATCA 2727
Db
966 LeuProValProValHisTyrAlaHisLeuSerCysGlu-----Lys 979
QY
2728 GTCTGTGCTCCCATCTGTACGCACATCTGGCAGCTGCTCAGGTTGGCCAGTTTCATAAAG 2787
Db
980 AlaLysGluLeuTyrArgThrTyrLysGluHis 990
QY
2788 TTCGATGAGATGTCGGAGAGCTCTCTCCAGTCAT 2820
Db

RESULT 13

US-11-218-305-14683
; Sequence 14683, Application US/11218305
; Publication No. US2006014195A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14683
; LENGTH: 2012
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-14683

Alignment Scores:

Pred. No.: 1,17e-37 Length: 2012
Score: 466.00 Matches: 166
Percent Similarity: 42.6% Conservative: 97
Best Local Similarity: 26.9% Mismatches: 229
Query Match: 8.7% Indels: 126
DB: 7 Gaps: 24

US-10-645-746-3 (1-1020) x US-11-218-305-14683 (1-2012)

QY 392 IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTyrAlaVal 411
Db 27 ATTCAGCATCTCGCAGCAGAGACCATTTACCTAGAACATCTTAATCGAAAAACGAACACTATTTA 86
QY 412 ProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIleAspLeu 431
Db

Db 87 CCAATGGAGCGCTGCAAGATCGTTGAA---GGCCAGAGATACACGAAGAGTTGTAATGAA 143
Qy 432 ValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLysGluAsn 451
Db 144 AAACAGATCACATCGTTGCTAAAGGTTACATGTCMAAGGCTCGAGAACAGAGATGGAT 203
Qy 452 ThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheValGluArg 471
Db 204 ATTCTACACAGACATTCATCAAAATGGATAT-----GAGCAAGATCCATATCGAAGGAA 257
Qy 472 PheGly-----LeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
Db 258 TTTGGGATCAACATTAGTCAGAAGCTAACCTATGTTGAA-----GCCGAGTCCCTTCCT 311
Qy 490 GluProMetLeu---ValAsnSerValAsnGluGlnIleLysMetThrProValIleArg 508
Db 312 GCACCTTGCTGAAGTATCATGACACTGGAAAGAGAAAGAGTGTCTTACCA-----362
Qy 509 GlyPheGlnGluLysGlnLeuAsnValValProGluLysGluLeu---CysCysAlaVal 527
Db 363 -----CAGTTGCTCAGTGGACATGGTAACAAGAAAGATGTCAAAATTCGGGCATGGAAATTTAACAGT 527
Qy 565 AsnGluAsnArgGlyAlaGlnSerIleMetTyrAsp-----AlaThrLysAsn 580
Db 528 GAGCCGCGTATGCCAATATATTACGCTAGACACCATCAAGTAGTGAAGGCACCTTAAAGT 587
Qy 581 GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAla 600
Db 588 GTGTAT-----AATATTGCACCTGAAC-----608
Qy 601 AlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMet 620
Db 609 -----AAACTCAAGGGTAAAGAACTTGAACCTCTCTCTG 641
Qy 621 PheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAsp 640
Db 642 ---GCTATACTCCCGACACAACTGGTCCGTATATATGTTGATCATCAAGATATTTGTGAA 698
Qy 641 HisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAla 660
Db 699 ACTGATTTGGGATTGATATACAAATGTTGCTTAAACCAAGCATGTTTAAAGATC-----752
Qy 661 SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsn 680
Db 753 -----AGCAAAACAGTACTTGGCAAAATGTCTCACTGAAAAATTAAT 791
Qy 681 AlaLysLeuGlyGlyIleAsn-----GlnGluLeuAspTyrSerGluIleAla 696
Db 792 GTTAGATGGGAGGAGAAACACTGTGCTCGGACGCAATAAGTTGGAGC-----842
Qy 697 GluIleSerProGluLysGluArgLysThrMetProLeu-----711
Db 843 -----ATTCCTTTGGTCAGTGACATCCCA 866
Qy 712 ThrMetTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSer 731
Db 867 ACTATATATTTGGTGACATGTAAACACACCTCGAAACCGGGAGGAGCTCAAGTCCATCA 926
Qy 732 IleAlaAlaValValAlaSerIleAsn---ProGlyGlyThrIleTyrArgAsnMetIle 750
Db 927 ATCGCTGCGGTTGCTTCTCAAGATTTGGCCAGAGTTACAAAGTATGCTGATGTTGGT 986
Qy 751 ValThrGln-----GluGluCysArgProGlyGluArgAlaValAlaHisGlyArg 767
Db 987 TGTGCTCGGCACACCGGCAAGAGCTCAATTCAGGACCTTTACAAACCATGGCACGATCCT 1046

Qy 768 GluArgThrAspIleuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsn 787
Db 1047 CAGAGAGCAGCTGTAAACAGCGCGCATGATCAGGGAGCTGTTAATATCTTCAGGAAGGCC 1106
Qy 788 AsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGlu 807
Db 1107 ACTGGGCAGAG---CCATTGAGATATATTTCTACAGGACGGTGTATTAGTGAAGCCAG 1163
Qy 808 MetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMet 827
Db 1164 TTCTATCAAGTCTCTCTTACGAGTTAGTGCATCCGTAAGGCATCGCATCCCTA---1220
Qy 828 SerGluArgAspGlyGluAspProGluProLysTyr-----ThrPheIleVal 843
Db 1221 -----GAACCAAAATTACCAGCTCTCTGTAAACATTTGGGTG 1256
Qy 844 IleGlnLysArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValVal 863
Db 1257 GTTCAAAACGTCATCATACGAGACTATTTACAAACATCACAAGACAGAGTAGCATG 1316
Qy 864 AsnLysAspLeuThrProAlaGluThrAspValAlaValAlaAlaValLysGlnTrpGlu 883
Db 1317 GACAAG-----1322
Qy 884 GluAspMetLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrVal 903
Db 1323 -----AGTGAATAATTTTGCCA-----GGAAGTGTGT 1352
Qy 904 AspLysLeuIleValSerLysTyrLysPheAspPhePheLeuAlaSerHisGlyVal 923
Db 1353 GATTCTAAGATATGCCACCCACACAGATTGATTCTACCTCTGTAGTCATGCTGGAATC 1412
Qy 924 LeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGln 943
Db 1413 CAGGAAACAAGTAGGCGGCTCACTACCATGTCTCTGGGATGAGAACAAATTTTCACAGCA 1472
Qy 944 AspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysPro 963
Db 1473 GACGAATGCAAAACACTGCACAAACACCTTGTCTACATTATGCCGTTGCACACGCTCG 1532
Qy 964 IleSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAlaLys 981
Db 1533 GTTCTGTGTTGCTCTCTGCTACTACTACGACACCTGGCAGCATTCGGGCGCGG 1586

RESULT 14

US-11-218-305-19467
; Sequence 19467, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLeod, Paul L.
; APPLICANT: Tao, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19467
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-19467

Alignment Scores:
Pred. No.: 1,976-36 Length: 2773
Score: 456.00 Matches: 204
Percent Similarity: 39.9% Conservative: 125

Best Local Similarity:	24.7%	Mismatches:	313
Query Match:	8.5%	Indels:	183
DB:	7	Gaps:	35
US-10-645-746-3 (1-1020) x US-11-218-305-19467 (1-2773)			
QY	225	ValProGluSerPheHisAspProAsnArgPheGluGlnSerLeuGluValAlaProArg	244
DB	341	GTGGTAGGTCATTTTATCTCCAC---TTAGGGAGACCCCCCCCCCCCCCCT	397
QY	245	IleGluAlaTrpPheGlyIleTyrlleGlyIleGlyLeuPheAspGlyGluProVal	264
DB	398	TTGGAAAGTTGGCGCGTTTCTACAAAGCATAGGCGGCACAGATGGGCTTCA	454
QY	265	LeuAsnPheAlaIleValAspLeuPheTyran---AlaProLysMetSerLeuLeu	283
DB	455	CTGAATATTGATATGCTCTACTGCAATTTATCGAGCCTCTCCCTGATTTGTT	514
QY	284	AspTyrlleLeuLeu-----LeuIleValAspProGlnSerCysAsnAspValArg	300
DB	515	GCTCAGCTTCTTAAGAGATATTTAGTTAGGCATTTGCTGATTTCTGATCGCGTGAAG	574
QY	301	LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro	320
DB	575	-----ATCAAAAAGCCTTAAGAGGTGTGAGGTTGAGGTCTACTCACAGGGAAACATG	628
QY	321	ArgIleArgGlnLeuLeuGluAsnLeuLysCysAlaGluValTrpAspAsnGlu	340
DB	629	CGCAGAAAGTATCGCATTTCTGCGCTC-----	655
QY	341	MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeu	360
DB	656	ACCTCAACAGCAACA---AGAGAGCTATCATTC-----	685
QY	361	ValTyrlsValThrGlyLysSerAspArgGlyArgAsnAlaLysTyrlsAspThrThr	380
DB	686	-----CCTGTTGATGATCGTGGTACT-----GTGAGACT	715
QY	381	LeuPheLysIleTyrlleGluGluAsnLysLysPhe-----IleGluPheProHisLeuProLeu	399
DB	716	GTGGTCCCAATCTTCAATGAGACTTATGGTTTATGATCCAGCACACCACTTTACCGTGC	775
QY	400	ValLysValLysSerGlyAlaLysGluTyrlAlaValProMetGluHisLeuGluValHis	419
DB	776	TTACAAGTGGCAATCAACAAAGACCCAAATTATCTGCTATGAGGTTTGCAGATAGTT	835
QY	420	GluLysProGlnArgTyrlsAsnArgIleAspLeuValMetGlnAspLysPheLeuLys	439
DB	836	GAA---GGACAGCTTACTCAAGCGACTCAATGAGAAACAAATCACTGCTCTACTGAA	892
QY	440	ArgAlaThrArgLysProHisAspTyrlsGluAsnThrLeuLysMetLeuLysGluLeu	459
DB	893	GTGACCTGCGCGCCCT-----CAAGACGTGAGCTGGACATTTTACAGACTGTG	943
QY	460	AspPheSerSer---GluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeu	478
DB	944	CATCAATCGCTACTATGAGACCCGTATGCAAGAAATTTGGTATAGAAATTTGATGAA	1003
QY	479	GlnMetIleGluCysProGlyLysValLeuLysGluProMetLeu---ValAsnSerVal	497
DB	1004	CGCCTTGCTGAGTGAAGCTCGTGTCTGCCACCAAGGCTTAAATACCATGATAGT	1063
QY	498	AsnGluGlnIleLysMetThrProValIle-----ArgGlyPheGlnGluLysGln	514
DB	1064	GGCCGAGAGAGGATGTTTGGCCAGATTTGGCCAATGGAACATGATGAATAAGAAATG	1123
QY	515	LeuAsnValValProGluLysGluLeuCysCysAlaValPheValValAsnGluThrAla	534
DB	1124	GTAAATGGGAGATCAGCAACTGGCGATGATTAACTTCTCTCGGAAT-----	1174
QY	535	GlyAsnProCysLeuGluGluAsnAspValValLysPheTyrlThrGluLeuIleGlyGly	554
DB	1175	-----GTGCAAGATAGTGGCGGTCTGTGTCATGCAACTGGCAATCATG	1222
QY	555	CysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMet	574
DB	1223	TGCCAAATATACAGGAATG-----GATTTTCTCTTGAGCCTGTGCTG	1264
QY	575	TyrAspAlaThrLysAsn-----GluTyrlAlaPheTyrlLysAsnCysThrLeu	590
DB	1265	CTCCAGTGTGATGCAAGCCAGAACATGTTTGAAGAGCGTTGAAG-----	1309
QY	591	AsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArg	610
DB	1310	-----GCAGTTATCAA-----GATGCATGAACATACTAGG---	1342
QY	611	LeuProAspLysGluGlnLysValLeuMetPheIleIleLysSerLysArgGlnLeuAsn	630
DB	1343	---CCACAGGGAGGAACTTGATCTGCTGATTTAATATCTGCTGATTAATGTTTCC	1399
QY	631	AlaTyrlGlyPheValLysHisTyrlCysAspHisThrIleGlyValAlaAsnGlnHisIle	650
DB	1400	TTATATGGGATCTCAAAAGGATCTGTGAGACTGTGATCTCGGATTTGGTCTCCCATGTTGT	1459
QY	651	ThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArg	670
DB	1460	CTGACTAAACATGTTTTTAAGATG-----AGCAGCAG	1492
QY	671	IlePheTyrlGlnIleAlaLeuLysIleAsnAlaLysLeuGly---GlyIleAsnGlnGlu	689
DB	1493	TATCTTGCAAAATGTTGCACTCAAAATAAATGTTAAGTTGGCGGGCGCCCGCCT	1552
QY	690	LeuAspTrpSerGluIleAlaGluIleSerProGluGluLysGluArgArgLysThrMet	709
DB	1553	TTAGAT-----GCTTTGACAAGAGGAATC	1576
QY	710	ProLeu-----ThrMetTyrlValGlyIleAspValThrHisProThrSer	724
DB	1577	CCCTTGTGTCAGTCAGACAGCCGACCAATAATTTGCTGCTGATGTTACCATCCATCCT	1636
QY	725	TyrSerGlyIleAspTyrlSerIleAlaValAlaSerIleAsn-----ProGlyGly	743
DB	1637	GGAGAAGATTCACAGTCTCTCCATTGCACTGTGTTGCTTTCGCAAGACTGGCTGAGTC	1696
QY	744	ThrIleTyrlArgAsnMetIleValThrGlnGluCysArgProGlyGluArgAlaVal	763
DB	1697	ACCAATATGCTCGACTAGTAGTGCCCAA-----	1726
QY	764	AlaHisGlyArgGlu-----ArgThr	770
DB	1727	GCCATCGCCAGGAGCTGATACAGGATCTTTCAAAGTATGGCAAGATCCACAGAGAAG	1786
QY	771	AspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsn	790
DB	1787	ACAGTAACTGGTGGCATGATAAGGAATCTCTCATTTCTTTC---AAGAGAGCAACTGGA	1843
QY	791	ArgAlaProAlaHisIleValValTyrlArgAspGlyValSerAspSerGluMetLeuArg	810
DB	1844	CAGAAGCCCCAGAGGATCATATTTCTACAGGATGGTGTGAGTGGAGACATCTTATCAA	1903
QY	811	ValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArg	830
DB	1904	GTATTGTTGATGAATCTGATGCCATGCAAGGCTGTGTCATCTTG-----	1951
QY	831	AspGlyGluAspProGluProLysTyrl-----ThrPheIleValIleGlnLys	846
DB	1952	-----GAGCCCAACTACCAAGCTCCAGTCTTCTTCTGCTGGTGCAGAAA	1996
QY	847	ArgHisAsnThrArgLeuLeuArgMetGluLysAspLysProValValAsnLysAsp	866
DB	1997	CGACATCACACATGAGTGTGTTTCTAATAACCAACAGATCAGGTACAGTTGATAGA---	2053
QY	867	LeuThrProAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMet	886
DB	2053	-----	2053

```
QY 887 LysGluSerLysGluThrGlyIleValAenProSerSerGlyThrValAspLysLeu 906
Db 2054 -----AGCGGAACATACATGCT-----GGCACCGTGTGATTGGAAG 2092
QY 907 IleValSerLysThrLysPhePheLeuAlaSerHisGlyValLeuGlyThr 926
Db 2093 ATTTGCCATCTACTGAATTTGTTCTACTGTGTAGCCATCTGGCATTGAGGAC 2152
QY 927 SerArgProGlyHisThrValMetTyrAspAspLysGlyMetSerGlnAspGluVal 946
Db 2153 AGCGCGCTCTCAATTACCATCTCTGTGGGACGAGAACAGTTCCAGCTGATGCTG 2212
QY 947 TyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaAaCysArgLysProIleSerLeu 966
Db 2213 CAGACTCTGACAAACACTATGCTACAGTACGTAGTGTGACCGCTCCGTGCAATT 2272
QY 967 ProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThr 986
Db 2273 GTGCCCCCGCATCATCTGCTGCGGCGCTTCCGAGCT-----CGCTTC 2320
QY 987 TyrLysGluHisTyrIleGlyAsp-TyrAlaGln-----ProArgThrArgHis 1002
Db 2321 TACATGGACCCAGATACCTCTGACATGGCTCAATGGCCAGTGGTCCCGTCCCTCCA 2380
QY 1002 eGluMetGluHis 1006
Db 2381 CCAGGTGGCGCAC 2393

RESULT 15
; US-11-218-305-19466
; Sequence 19466, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19466
; LENGTH: 3827
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (480)..(480)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (515)..(515)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (568)..(568)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (649)..(649)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (787)..(787)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (964)..(970)
```

```
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-19466
Alignment Scores:
Pred. No.: 3,97e-36 Length: 3827
Score: 455.00 Matches: 196
Percent Similarity: 40.6% Conservative: 126
Best Local Similarity: 24.7% Mismatches: 301
Query Match: 8.5% Indels: 173
DB: 7 Gaps: 32

US-10-645-746-3 (1-1020) x US-11-218-305-19466 (1-3827)
QY 225 ValProGluSerPheHisAspProAenArgPheGluGlnSerLeuGluValAlaProArg 244
Db 1388 GTCGGTAGTCTATTTTACTCTCCCAAC---TTAGGAGAGAGCTCAAAAACCTTGGTAGGGA 1444
QY 245 IleGluAlaTTPheGlyIleTyrIleGlyIleLysGluLeuPheAspGlyGluProVal 264
Db 1445 TTGGAAAGTTGGCGTGGTTTTTACCAAGCATAGCCGACACA-GATGGGCTTT-TCA 1502
QY 265 LeuAenPheAlaIleValAspLysLeuPheTyrAen---AlaProLysMetSerLeuLeu 283
Db 1503 CTGAATATTGATATGCTCTACTGCAITTTATCGAGCCTCTCCTGTGATCGATTTGTT 1562
QY 284 AspTyrLeuLeu-----LeuIleValAspProGlnSerCysAenAspAspValArg 300
Db 1563 GCTCAGCTTCTTAACAGAGATATCTCAGTAGGCCATTTCTGTGATCTGCGTGAAG 1622
QY 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro 320
Db 1623 -----ATTAAAAAGCCCTAAGAGGTGTGAAGTTGAGGTGACTCACAGGGGAAACATG 1676
QY 321 ArgIleArgGlnLeuLeuGluAsnLysLysLeuLysCysAlaGluValTrpAspAenGlu 340
Db 1677 CGCAGAAAATATCGCATTTCTGGCCTC-----1703
QY 341 MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuLysGluGluAenSerLeu 360
Db 1704 ACCTCACAGCAACA---AGAGAGCTATCATTC-----1733
QY 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThr 380
Db 1734 -----CCTGTTGATGATCGTGTACT-----GTGAAGACT 1763
QY 381 LeuPheLysIleTyrGluGluAsnLysLysPhe---IleGluPheProHisLeuProLeu 399
Db 1764 GTGGTGCAATACTTTCATGGAGACTTATGGTTTTTAGTATCCACACACCATCTTACCATGC 1823
QY 400 VallysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluHis 419
Db 1824 TTGCAAGTGGGTAATCAACAAGACCAAAATATCTGCTTATGGAGGTTTGCAGATAGATT 1883
QY 420 GluLysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLys 439
Db 1884 GAA---GGCAGCGTTACTCAAGCGACTCAATGAGAAAACAATCACTCTCTACTGAAA 1940
QY 440 ArgAlaThrArgLysProHisAspTyrLysGluAenThrLeuLysMetLeuLysGluLeu 459
Db 1941 GTGACCTGCCAGCGCCT-----CAAGAGCGGAGCTGGACATCTTACAGACTGTG 1991
QY 460 AspPheSerSer---GluGluLeuAenPheValGluArgPheGlyLeuCysSerLysLeu 478
Db 1992 CATCAATGATCATATGTAAGACCCCTATGCACTGGAAATTTGGTATAGAAATGATGAA 2051
QY 479 GlnMetIleGluCysProGlyLysValLeuLysGluProMetLeu---ValAenSerVal 497
Db 2052 CGTCTGTGTCAGTTGAAGCTCGTGTCTGCCACCACCAAGACTTAAATACCATGATAGT 2111
QY 498 AsnGluGlnIleLysMetThrProValIle-----ArgGlyPheGlnGluLysGln 514
Db 2112 GCGCCGAGAGAGGATGTTTTCGCCCACTGGCCCAATGGAAACATGATGAATAAGAAATG 2171
```

```
Qy 515 LeuAsnValValProGluLysGluLeuCysCysAlaValPheValValAsnGluThrAla 534
Db : : : : :
2172 GTTAATGTTGCGACAGTGGCACTGGGCATGTATTAACTTCTCTCGGAAT- 2222
Qy 535 GlyAsnProCysLeuGluGluAsnAspValValLysPheTyrThrGluLeuIleGlyGly 554
Db : : : : :
2223 -----GTGCAAGATAGTGGCCCTAGGCGTTCTCTCATGAGTTGGCATCATG 2270
Qy 555 CysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMet 574
Db : : : : :
2271 TGCCTGCAATATCAGGAATG-----GATTTTGTCTTGTAGCCTGTGCTG 2312
Qy 575 TyrAspAlaThrLysAsn-----GluTyrAlaPheTyrLysAsnCysThrLeu 590
Db : : : : :
2313 CCTCAGTACTGCAAGCCAGAACATGTTGAGAGAGCGTTAAAG----- 2357
Qy 591 AsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArg 610
Db : : : : :
2358 -----GCACGTTATCAA-----GATGCAATGAACATACTGAGG--- 2390
Qy 611 LeuProAspLysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsn 630
Db : : : : :
2391 ---CCACAGGAGGAGGAATGATCTGCTGATCGTAATACTGCTGACCAACAATGTTCT 2447
Qy 631 AlaTyrGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAlaAsnGlnHisIle 650
Db : : : : :
2448 CTTTATGGGATCTCAAAAGATCTGTGAGACTGAATCGGATGGTCTCCCACTGTTGT 2507
Qy 651 ThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArg 670
Db : : : : :
2508 CTGACTAAACATGTTTTAGATG-----AGGCAACAG 2540
Qy 671 IlePheTyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeu 690
Db : : : : :
2541 TACCTTGCAAAATGTTGCACTCAAAATAATGTTAAGGTGGGGAAGGAATACTGTACTT 2600
Qy 691 AspTrpSerGluIleAlaGluIleSerProGluLysGluArgArgLysThrMetPro 710
Db : : : : :
2601 -----TTAGATGCTTTGTCTCA-----AGGAGAATCCCC 2627
Qy 711 Leu-----ThrMetTyrValGlyIleAspValThrHisProThrSerTyr 725
Db : : : : :
2628 CTTGTCACTGACAGACCGACCAATAATTTGGTCTGATGTTACCCATCCATCCTGGA 2687
Qy 726 SerGlyIleAspTyrSerIleAlaAlaValAlaSerIleAsn---ProGlyGlyThr 744
Db : : : : :
2688 GAAGATTCCAGTCTCTCCATTGCGCCGTTGTTGCTTCGCAAGACTGCGCCGAGTCAAG 2747
Qy 745 IleTyrArgAsnMetIleValThrGlnGluCysArgProGlyGluArgAlaAla 764
Db : : : : :
2748 AATACGCTGGACTAGTAGTGGCAA-----GCC 2777
Qy 765 HisGlyArgGlu-----ArgThrAsp 771
Db : : : : :
2778 CATGCCGAGGAGCTGATACAGGATCTTTCAAAGATAGCGAGCCGCGAGAGGAGCG 2837
Qy 772 IleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArg 791
Db : : : : :
2838 GTAACCTGGCGCATGATAAAGGAATCTCTATTTCTTC---AAGAGGGCAACTGGACAG 2894
Qy 792 AlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgVal 811
Db : : : : :
2895 AAGCCCCAGAGGATCATATTCTACAGGATGGTGTCTAGTGGGACAGATTTCTATCAAGTA 2954
Qy 812 SerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAsp 831
Db : : : : :
2955 TTGCTGTACGAATTTGATGTCATTAGAAAGGCTGTGCGTCCCTG----- 2999
Qy 832 GlyGluAspProGluProLysTyr-----ThrPheIleValIleGlnLysArg 847
Db : : : : :
3000 -----GAGCCCAACTACCAGCTCCAGTTACTTTTGTCTGTTGTTACAGAGCGC 3047
Qy 848 HisAsnThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeu 867
```

```
Db : : : : :
3048 CATCACACTAGGCTGTTTGGAAACAACACACAGTATCAGCGCACAGTCGATAGA----- 3101
Qy 868 ThrProAlaGluThrAspValAlaAlaValAlaValLysGlnTrpGluGluAspMetLys 887
Db : : : : :
3101 ----- 3101
Qy 888 GluSerLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIle 907
Db : : : : :
3102 -----AGCGGAACAACATACTGCT-----GGCACCGTGGTTCGATTTCGAAGATT 3143
Qy 908 ValSerLysTyrLysPheAspPhePheLeuAlaSerHisHisGlyValLeuGlyThrSer 927
Db : : : : :
3144 TGCATCCTACTGAGTTTGACTTCTTACTGTGTGATCCATGCTGCATTTACGGGAACGAGC 3203
Qy 928 ArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyr 947
Db : : : : :
3204 CGCCTGCTCACTACCATGCTGCTGTGGAGCGAGAACAAAGTTCACAGCTGACAGCTGCAG 3263
Qy 948 LysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuPro 967
Db : : : : :
3264 ACCCTGACGAACAACCTGTGCTACACGTACGCTAGGTGACCCCGCTCGTGTCCATCGTG 3323
Qy 968 ValProValHisTyrAlaHisLeuSerCysGluLysAlaLys 981
Db : : : : :
3324 CCCCCGCGTACTACGCTCATCTGGCAGCCTTCCGAGCTCGC 3365
```

Search completed: July 6, 2006, 07:16:19
Job time : 1546 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 21:39:41 ; Search time 14570 Seconds
(without alignments)
12385.153 Million cell updates/sec

Title: US-10-645-746-2
Perfect score: 3227
Sequence: 1 cagccacaagtgatgaac.....tttaaaaaaaaaaaaaaa 3227

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	782.4	24.2	796	2	BJ155598
C 2	774	24.0	786	2	BJ143897
C 3	697.4	21.6	700	2	BJ150328
C 4	695.8	21.6	710	2	BJ132853
C 5	689	21.4	700	2	BJ155451
C 6	677	21.0	677	2	BJ779221
7	629.8	19.5	633	2	BJ122122
8	603.4	18.7	635	2	BJ755060
9	582	18.0	583	2	BJ116268
10	563.4	17.5	565	4	CB401772
11	555.4	17.2	557	2	BJ105221
12	554	17.2	554	4	CB398209
13	534	16.5	534	2	BJ127233
14	527	16.3	539	2	BJ127087
C 15	416.4	12.9	449	9	D27223
C 16	393.8	12.2	421	9	D27221
17	375	11.6	378	4	C62850
C 18	370.8	11.5	375	9	D32492
19	359	11.1	360	4	C60407

20	358	11.1	360	9	D35146
21	356.4	11.0	360	9	D35838
22	354	11.0	356	10	Z14900
23	348	10.8	360	4	C65335
24	345.4	10.7	370	1	AV191613
25	339.8	10.5	345	9	D33093
26	339.6	10.5	373	4	C60787
27	331	10.3	376	4	C60885
28	309	9.6	340	9	D27220
29	293.4	9.1	334	10	M89235
C 30	289.8	9.0	300	4	C51768
C 31	288	8.9	300	4	C51841
C 32	285	8.8	300	4	C51449
C 33	282.2	8.7	300	4	C53264
34	276.4	8.6	295	9	D27222
C 35	262	8.1	300	4	C54925
36	77	2.4	701	6	AV809756
C 37	77	2.4	701	9	CX861295
C 38	77	2.4	715	3	BU723459
39	71.8	2.2	834	13	CZ532832
40	69.2	2.1	749	8	CV511740
41	69.2	2.1	773	8	CV508440
42	69	2.1	577	2	BJ113168
C 43	69	2.1	623	9	DN303856
C 44	68.8	2.1	667	1	AU217048
C 45	68.8	2.1	703	2	BJ815785

ALIGNMENTS

RESULT 1
BJ155598/c
LOCUS
DEFINITION
796 bp mRNA linear EST 24-JAN-2002
Caenorhabditis elegans cDNA library, C. elegans L1 stage
ACCESSION
BJ155598
VERSION
BJ155598.1 GI:18323583
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 796)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C. elegans genome
Unpublished (2002)
JOURNAL
Contact: Tadasu Shin-i
COMMENT
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
source
1..796
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK134B04"
/sex="hermaphrodite"
/tissue type="whole animal"
/dev stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

ORIGIN

Query Match 24.2%; Score 782.4; DB 2; Length 796;
Best Local Similarity 99.6%; Pred. No 4e-174; 2; Indels 1; Gaps 1;
Matches 794; Conservative 0; Mismatches 0

Db	400	TCATGGTGTCC	TTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACGATRAAGG	341					
Qy	2840	AATGAGCCAGATGAAGCTCTATATAAATGACCTACGGACTTGCCTTTCTCTCTGTAGATG	2899						
Db	340	AATGAGCCAGATGAAGCTCTATATAAATGACCTACGGACTTGCCTTTCTCTCTGTAGATG	281						
Qy	2900	TCGAAAAACCCATCTCGTTGGCTGTTCGCGTTCATTATGCTCATTTATCATGTGAAAAAGC	2959						
Db	280	TCGAAAAACCCATCTCGTTGGCTGTTCGCGTTCATTATGCTCATTTATCATGTGAAAAAGC	221						
Qy	2960	GAAGAGCTTTATCGAACTTACAAAGNAATTATCATCGGTGACTATGCAAGCACCGGAC	3019						
Db	220	GAAGAGCTTTATCGAACTTACAAAGNAATTATCATCGGTGACTATGCAAGCACCGGAC	161						
Qy	3020	TCGACACGAATGGAACATTTCTCCAAACTGAAGTACCCCTGGAAATGCGTTCGC	3079						
Db	160	TCGACACGAATGGAACATTTCTCCAAACTGAAGTACCCCTGGAAATGCGTTCGC	101						
Qy	3080	ATAACATTTTGCAAAAGTGTGCGCCGTTTCAAATCAAATTTTCAAATTTGTAGATATTGTAC	3139						
Db	100	ATAACATTTTGCAAAAGTGTGCGCCGTTTCAAATCAAATTTTCAAATTTGTAGATATTGTAC	41						
Qy	3140	TTACTTTTTTTTAAAGCCCGGTTTCAAAAATTCATTTCCA	3178						
Db	40	TTACTTTTTTTTAAAGCCCGGTTTCAAAAATTCATTTCCA	2						
RESULT 4									
LOCUS	BJ132853/c								
DEFINITION	BJ132853 unpublished oligo-capped cDNA library, C. elegans LI stage	710 bp	mRNA	linear EST 23-JAN-2002					
ACCESSION	BJ132853								
VERSION	BJ132853.1	GI:18293010							
KEYWORDS	EST.								
SOURCE	Caenorhabditis elegans								
ORGANISM	Caenorhabditis elegans								
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.								
AUTHORS	Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y., and Sugano, S.								
TITLE	A complementary view of the C.elegans genome								
JOURNAL	Unpublished (2002)								
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.								
FEATURES	Location/Qualifiers								
source	1. .710								
	/organism="Caenorhabditis elegans"								
	/mol_type="mRNA"								
	/strain="N2"								
	/db_xref="taxon:6239"								
	/clone="yk1070f12"								
	/sex="hermaphrodite"								
	/tissue_type="whole animal"								
	/dev_stage="L1"								
	/clone_lib="unpublished oligo-capped cDNA library, C. elegans LI stage"								
ORIGIN									
Query Match	21.6%;	Score 695.8;	DB 2;	Length 710;					
Best Local Similarity	99.6%;	Pred. No. 1.3e-153;							
Matches 708;	Conservative	0;	Mismatches 2;	Indels 1; Gaps 1;					
Qy	2488	TAAAACAATTCATGTCGGAACCGGATGGAGAGATCCAGAGCCGAGTACAGTTCATTG	2547						
Db	710	TAAACCAATTCATGTCGGAACCGGATGGAGAGATCCAGAGCCGAGTACAGTTCATTG	651						


```
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 677)
AUTHORS
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
TITLE
A complementary view of the C.elegans genome
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..677
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1409d07"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

ORIGIN
Query Match 21.0%; Score 677; DB 2; Length 677;
Best Local Similarity 100.0%; Pred. No. 3.7e-149;
Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 ATGGAGAGATCCAGAGCCGAGTACACGTTCTATTGTGATTCGAAAGACACAATACAC 2571
DB 677 ATGGAGAGATCCAGAGCCGAGTACACGTTCTATTGTGATTCGAAAGACACAATACAC 618
QY 2572 GATTGCTTCCAGAGATCGAAAGATTAAGCCAGTGTCAATAAGATCTTACTCTGCTG 2631
DB 617 GATTGCTTCCAGAGATCGAAAGATTAAGCCAGTGTCAATAAGATCTTACTCTGCTG 558
QY 2632 AACACAGATGTCGTGTTGCTGTTTAAACAATGGGAGGAGGATATGAAAGAACAAAG 2691
DB 557 AACACAGATGTCGTGTTGCTGTTTAAACAATGGGAGGAGGATATGAAAGAACAAAG 498
QY 2692 AAATCGAATGTGAACCCATCATCCGGAACAACCTGTGGATAAATCTATCGTTTCGAAT 2751
DB 497 AAATCGAATGTGAACCCATCATCCGGAACAACCTGTGGATAAATCTATCGTTTCGAAT 438
QY 2752 ACAAAATTCGATTTTCTTGGCATCTCATCATGCTGCTTCTGTGTACATCTCGTCCAGGAC 2811
DB 437 ACAAAATTCGATTTTCTTGGCATCTCATCATGCTGCTTCTGTGTACATCTCGTCCAGGAC 378
QY 2812 ATTACACTGTTATGTATGACGATAAAGGAATGAGCCAGATGAAGTCTATAAATGACCT 2871
DB 377 ATTACACTGTTATGTATGACGATAAAGGAATGAGCCAGATGAAGTCTATAAATGACCT 318
QY 2872 ACGGACTGCTTTTCTCTGCTAGATGTGCAAAACCCATCTCGTTGCCCTGTTCCGGTTC 2931
DB 317 ACGGACTGCTTTTCTCTGCTAGATGTGCAAAACCCATCTCGTTGCCCTGTTCCGGTTC 258
QY 2932 ATTATGCTCATTTTATCATGTGAAAGCCGAAAGAGCTTTATCGAATTCAGAGAACATT 2991
DB 257 ATTATGCTCATTTTATCATGTGAAAGCCGAAAGAGCTTTATCGAATTCAGAGAACATT 198
QY 2992 ACATCGGTGACTATGCACAGCCAGCAGCTCGACACGAAATGGAACATTTTCTCCAACTA 3051
DB 197 ACATCGGTGACTATGCACAGCCAGCAGCTCGACACGAAATGGAACATTTTCTCCAACTA 138
QY 3052 ACGTGAAGTACCTCGGAATGTGCTTCGCAFAACATTTTGCAAAAGTGTGCCCGCTTCAA 3111
DB 137 ACGTGAAGTACCTCGGAATGTGCTTCGCAFAACATTTTGCAAAAGTGTGCCCGCTTCAA 78
QY 3112 TCAAAATTTTCAATTTAGATATTGTACTACTTTTAAAGCCCGGTTTCAAAATTT 3171
```

```
Db 77 TCAAAATTTTCAATTTAGATATTGTACTACTTTTAAAGCCCGGTTTCAAAATTT 18
QY 3172 CATTCATGACTTAACGT 3188
DB 17 CATTCATGACTTAACGT 1

RESULT 7
BJ122122
LOCUS
DEFINITION
BJ122122 unpublished oligo-capped cDNA library, C. elegans l1 stage
Caenorhabditis elegans cDNA clone yk1284h06 5', mRNA sequence.
ACCESSION
BJ122122
VERSION
BJ122122.1 GI:18282260
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 633)
AUTHORS
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
TITLE
A complementary view of the C.elegans genome
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..633
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1284h06"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="l1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans l1 stage"

ORIGIN
Query Match 19.5%; Score 629.8; DB 2; Length 633;
Best Local Similarity 99.7%; Pred. No. 5.7e-138;
Matches 631; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1377 CTAAAAATGCTGAAAGAATTTGGATTTCTCTTCTGAAAGAGCTAAATTTTGTGAAAGATTT 1436
DB 1 CTAAAAATGCTGAAAGAATTTGGATTTCTCTTCTGAAAGAGCTAAATTTTGTGAAAGATTT 60
QY 1437 GGATTTATGCTCCAACTTCAGATGATCGAATGTCAGGAAAGCTTTTGAAGAGCCAAATG 1496
DB 61 GGATTTATGCTCCAACTTCAGATGATCGAATGTCAGGAAAGCTTTTGAAGAGCCAAATG 120
QY 1497 CTTGTGAATGATGTAATGAACAAATTTAAATGACACCAAGTGTGCTGATTTCAAGAA 1556
DB 121 CTTGTGAATGATGTAATGAACAAATTTAAATGACACCAAGTGTGCTGATTTCAAGAA 180
QY 1557 AAACCAATTTGAATGTGTTCTCCGAAAAAGAACTTTGCTGTGCTGCTTTTGTAGTCAACGAA 1616
DB 181 AAACCAATTTGAATGTGTTCTCCGAAAAAGAACTTTGCTGTGCTGCTTTTGTAGTCAACGAA 240
QY 1617 ACAGCGGAAATTCATGCTTTAGAGAGAACGAGTTGTTAAAGTCTTACACCGAACTAATTT 1676
DB 241 ACAGCGGAAATTCATGCTTTAGAGAGAACGAGTTGTTAAAGTCTTACACCGAACTAATTT 300
QY 1677 GGTGTTGTCAAGTTCGTCGTAATACCAATTTGGTGGCAATGAAACAGAGAGGAGCGCAATCT 1736
DB 301 GGTGTTGTCAAGTTCGTCGTAATACCAATTTGGTGGCAATGAAACAGAGAGGAGCGCAATCT 360
```

```
QY 1737 ATTATGTACGACCGCGACGAAAAATGAATATGCTTCTTCAAAAAATGTACATTAATACC 1796
Db 361 ATTATGTACGACCGCGACGAAAAATGAATATGCTTCTTCAAAAAATGTACATTAATACC 420
QY 1797 GGAATCGGTAGATTTGAAATAGCCGCAACAGAGCGAAGAAATATGTTTGAAGCTCTTCCC 1856
Db 421 GGAATCGGTAGATTTGAAATAGCCGCAACAGAGCGAAGAAATATGTTTGAAGCTCTTCCC 480
QY 1857 GATAAGAACAAAAAGCTTTAATGTTTCATTATCATTTTCCAAACGCAACTGAATGCTTAC 1916
Db 481 GATAAGAACACAAAGCTTTAATGTTTCATTATCATTTTCCAAACGCAACTGAATGCTTAC 540
QY 1917 GGTTTTGTGAACATTTATGCGATCACACATCGGTGTAGCTTAATCAGCATATTAATCTT 1976
Db 541 GGTTTTGTGAACATTTATGCGATCACACATCGGTGTAGCTTAATCAGCATATTAATCTTCT 600
QY 1977 GAAACAGTCACAAAAGCTTTGGCATCACTAAGG 2009
Db 601 GAAACAGTCACAAAAGCTTTGGCATCACTAAGG 633

RESULT 8
BJ755060 635 bp mRNA linear EST 24-MAY-2004
LOCUS .
DEFINITION
  BJT55060 unpublished oligo-capped cDNA library Caenorhabditis
  elegans cDNA clone yk1409d07 5', mRNA sequence.
ACCESSION
  BJ755060
VERSION
  BJ755060.1 GI:47594822
KEYWORDS
  EST.
SOURCE
  Caenorhabditis elegans
  ORGANISM
    Caenorhabditis elegans
    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
    Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
  1 (bases 1 to 635)
AUTHORS
  Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
  and Sugano,S.
TITLE
  A complementary view of the C.elegans genome
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Tadaeu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
  Location/Qualifiers
    1..635
      /organism="Caenorhabditis elegans"
      /mol_type="mRNA"
      /strain="N2"
      /db_xref="taxon:6239"
      /clone="yk1409d07"
      /sex="Hermaphrodite"
      /tissue_type="whole animal"
      /dev_stage="varied"
      /clone_lib="unpublished oligo-capped cDNA library"

FEATURES
  source
    Query Match 18.7%; Score 603.4; DB 2; Length 635;
    Best Local Similarity 99.8%; Pred. No. 1e-131; 1; Indels 0; Gaps 0;
    Matches 604; Conservative 0; Mismatches 1;

QY 1 CAGCCCAAAAGTGATGAACATGCTCTCGAATTTTCCGAAATTTGGAAAAAGGATTTATC 60
Db 31 CAGCCCAAAAGTGATGAACATGCTCTCGAATTTTCCGAAATTTGGAAAAAGGATTTATC 90
QY 61 GTCATTTCTCGATCCGAGATGAATGCTTCGAGGCCCTCGTAATCGCAGCGCA 120
Db 91 GTCATTTCTCGATCCGAGATGAATGCTTCGAGGCCCTCGTAATCGCAGCGCA 150
QY 121 AATTCTTATGAGAAGAAAGTACTTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTT 180
```

```
Db 151 AATCTATGAGAGAAAGTACTTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTT 210
QY 181 ACGATCGGGAATACCTACGAGTATGAAGTGAATAATGACAAAGGAAGTATTTGAATAGAAAAC 240
Db 211 ACGATCGGGAATACCTACGAGTATGAAGTGAATAATGACAAAGGAAGTATTTGAATAGAAAAC 270
QY 241 CAGGAAAACCTTTCCCAAAAAGACAGAAATTTCCAATTTCCCGATCGTGCAAAACTCTTCT 300
Db 271 CAGGAAAACCTTTCCCAAAAAGACAGAAATTTCCAATTTCCCGATCGTGCAAAACTCTTCT 330
QY 301 GCGAACATCTTCGGCATGAGAGAAAGACAGACAGATTTTATTCTCGAAGACTATGTTTTTG 360
Db 331 GCGTACATCTTCGGCATGAGAGAAAGACAGACAGATTTTATTCTCGAAGACTATGTTTTTG 390
QY 361 ATGAAAAGGACACTGTTTATAGTGTTCGACTGAACTGACATCAAAATGCTGG 420
Db 391 ATGAAAAGGACACTGTTTATAGTGTTCGACTGAACTGACATCAAAATGCTGG 450
QY 421 TTTCCGAGAAAGTAGTAAAGAGGATTCGGAGAAAAGAGATGAAAGGATTTTGGAGAAA 480
Db 451 TTTCCGAGAAAGTAGTAAAGAGGATTCGGAGAAAAGAGATGAAAGGATTTTGGAGAAA 510
QY 481 AATCTTATACAAATGATCTTACCTATCGTAAAAAATTTTCCCTGAACTTTAGTCGAG 540
Db 511 AATCTTATACAAATGATCTTACCTATCGTAAAAAATTTTCCCTGAACTTTAGTCGAG 570
QY 541 AAAATCCGAAAAAGACGAAAGCGAATCGGAGTTACAAATTCCTGAAGATGTTATGA 600
Db 571 AAAATCCGAAAAAGACGAAAGCGAATCGGAGTTACAAATTCCTGAAGATGTTATGA 630
QY 601 CCCAG 605
Db 631 CCCAG 635

RESULT 9
BJ116268 583 bp mRNA linear EST 23-JAN-2002
LOCUS .
DEFINITION
  BJ116268 unpublished oligo-capped cDNA library, C. elegans L1 stage
  Caenorhabditis elegans cDNA clone yk1201c08 5', mRNA sequence.
ACCESSION
  BJ116268
VERSION
  BJ116268.1 GI:18276379
KEYWORDS
  EST.
SOURCE
  Caenorhabditis elegans
  ORGANISM
    Caenorhabditis elegans
    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
    Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
  1 (bases 1 to 583)
AUTHORS
  Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
  and Sugano,S.
TITLE
  A complementary view of the C.elegans genome
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Tadaeu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
  Location/Qualifiers
    1..583
      /organism="Caenorhabditis elegans"
      /mol_type="mRNA"
      /strain="N2"
      /db_xref="taxon:6239"
      /clone="yk1201c08"
      /sex="hermaphrodite"
      /tissue_type="whole animal"
      /dev_stage="L1"
      /clone_lib="unpublished oligo-capped cDNA library, C.
      elegans L1 stage"

FEATURES
  source
    1..583
      /organism="Caenorhabditis elegans"
      /mol_type="mRNA"
      /strain="N2"
      /db_xref="taxon:6239"
      /clone="yk1201c08"
      /sex="hermaphrodite"
      /tissue_type="whole animal"
      /dev_stage="L1"
      /clone_lib="unpublished oligo-capped cDNA library, C.
      elegans L1 stage"

ORIGIN
```


Db	374	AGAAATCGGAATTGTGAACCCATCATCCGGAACAACCTGTGGATAAACTTATCGTTTCGAA	315
Qy	2750	ATACAAATTCGAATTTTCTTTGGCATCTCATCATGGTGTCTTGGTACATCTCGTCCAGG	2809
Db	314	ATACAAATTCGAATTTTCTTTGGCATCTCATCATGGTGTCTTGGTACATCTCGTCCAGG	255
Qy	2810	ACATTACACTGTATTATGTATGACGATAAAGGAATGAGCCAGATCAAGTCTATAAAATGAC	2869
Db	254	ACATTACACTGTATTATGTATGACGATAAAGGAATGAGCCAGATCAAGTCTATAAAATGAC	195
Qy	2870	CTACGGACTTGCCTTTTCTCTCTGCTAGATGTCGAAAAACCCATCTCGTGTGCTGTTCGGT	2929
Db	194	CTACGGACTTGCCTTTTCTCTCTGCTAGATGTCGAAAAACCCATCTCGTGTGCTGTTCGGT	135
Qy	2930	TCATTATGCTCATTTATTCATGTGAAAAAGCGAAAGAGCTTTATCGAACTTTACAGGAACA	2989
Db	134	TCATTATGCTCATTTATTCATGTGAAAAAGCGAAAGAGCTTTATCGAACTTTACAGGAACA	75
Qy	2990	TTACATCGGTGACTATGACACAGCCACGGACTTCGACAGAAATGGAACATTTTCTCCAAC	3049
Db	74	TTACATCGGTGACTATGACACAGCCACGGACTTCGACAGAAATGGAACATTTTCTCCAAC	15
Qy	3050	TAACGTGAAGTACC	3063
Db	14	TAACGTGAAGTACC	1

RESULT 13	BJ127233	LOCUS	BJ127233	534 bp	mRNA	linear	EST 23-JAN-2002
DEFINITION	BJ127233 unpublished oligo-capped cDNA library, <i>C. elegans</i> L1 stage <i>Caenorhabditis elegans</i> cDNA clone yk1344b04 5', mRNA sequence.						

VERSION	BJL127233.1	GI:18287390
KEYWORDS	EST.	
SOURCE	Caenorhabditis elegans	
ORGANISM	Caenorhabditis elegans	
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;	
	Rhabditoides; Rhabditidae; Peloderinae; Caenorhabditis.	
REFERENCE	1. (bases 1 to 534)	
AUTHORS	Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.	
TITLE	A complementary view of the C.elegans genome	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Tadasu Shin-i Center for Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855	

```

FEATURES
source
1. .534
Location/Qualifiers
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1344b04"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cdna library, C.
elegans L1 stage"

```

Query Match	16.5%	Score 534;	DB 2;	Length 534;
Best Local Similarity	100.0%;	Pred. No. 2.8e-115;		
Matches 534;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1105	ATAAGTCACTGTTAAATCGCAGAGGAAGAAATGCAAAAAGTACGATCTACATTGT	1164	
Db	1	ATAAGTCACTGTTAAATCGCAGAGGAAGAAATGCAAAAAGTACGATCTACATTGT	60	

Qy	1165	TCAAAATCTATGAGGAAAAACAAAAGATTCTCATGTGATTTCCCAACCTTACCACTAGTCAAAG	1224
Db	61	TCAAAATCTATGAGGAAAAACAAAAGATTCTCATGTGATTTCCCAACCTTACCACTAGTCAAAG	120
Qy	1225	TTAAAGTGGAGCAAAAGAAATAGCTGTACCAATGGAAACATCTTTGAAGTTTCATGAGAAGC	1284
Db	121	TTAAAGTGGAGCAAAAGAAATAGCTGTACCAATGGAAACATCTTTGAAGTTTCATGAGNAGC	180
Qy	1285	CACAAAGATACAGAATTCGAATTTGATCTGTGTGATGCAAGAACAAGTTTCTTAAACGCGACTA	1344
Db	181	CACAAAGATACAGAATTCGAATTTGATCTGTGTGATGCAAGAACAAGTTTCTTAAACGCGACTA	240
Qy	1345	CACGAAACCTCAGCACTACAAAGAAATACCCCTTAAAAATGCTGGAAGAAATTCGATTTCT	1404
Db	241	CACGAAACCTCAGCACTACAAAGAAATACCCCTTAAAAATGCTGGAAGAAATTCGATTTCT	300
Qy	1405	CTTCTGAAGAGCTAAATTTTGTGTTGAAAGATTTTGGATTTATGCTCTCAAACTTTCAGATGATCG	1464
Db	301	CTTCTGAAGAGCTAAATTTTGTGTTGAAAGATTTTGGATTTATGCTCTCAAACTTTCAGATGATCG	360
Qy	1465	AAATGTCAGGAAAGGTTTTGAAAGAGCCAATGCTTGTGTAATAGTGTAAATGAACAAATTA	1524
Db	361	AAATGTCAGGAAAGGTTTTGAAAGAGCCAATGCTTGTGTAATAGTGTAAATGAACAAATTA	420
Qy	1525	AAATGACACACAGTGTATTCGTGGAGTTTCAAGAAAAACAATTGAATGTGGTTTCCCGAAAAAG	1584
Db	421	AAATGACACACAGTGTATTCGTGGAGTTTCAAGAAAAACAATTGAATGTGGTTTCCCGAAAAAG	480
Qy	1585	AACTTTGTCTGTCTGTTTTTTTGTAGTCAACGAAACAGCGGGAATTCATGCTTTAG	1638
Db	481	AACTTTGTCTGTCTGTTTTTTTGTAGTCAACGAAACAGCGGGAATTCATGCTTTAG	534

RESULT 14		BJ127087	539 bp	mRNA	linear	EST 23-JAN-2002
LOCUS		BJ127087	unpublished oligo-capped cDNA library,	C.elegans L1 stage		
DEFINITION		BJ127087	Caenorhabditis elegans cDNA clone yk1342dl1 5'	, mRNA sequence.		
ACCESSION		BJ127087				
VERSION		BJ127087.1	GI:18287244			
KEYWORDS		EST.				
SOURCE		Caenorhabditis elegans				
ORGANISM		Eukaryota; Metazoa; Nematoda; Chromadorea;	Rhabditida;			
		Rhabditoidea; Rhabditiidae; Peloderinae;	Caenorhabditis.			
REFERENCE		1 (bases 1 to 539)				
AUTHORS		Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.				
TITLE		A complementary view of the C.elegans genome				
JOURNAL		Unpublished (2002)				
COMMENT		Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.				
FEATURES		Location/Qualifiers				
source		1..539	/organism="Caenorhabditis elegans"			
			/mol_type="mRNA"			
			/strain="N2"			
			/db_xref="taxon:6239"			
			/clone="yk1342dl1"			
			/sex="hermaphrodite"			
			/tissue type="whole animal"			
			/dev_stage="L1"			
			/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"			
ORIGIN						
Query Match		16.3%	Score 527:	DB 2:	Length 539:	

Best Local Similarity 99.6%; Pred. No. 1.3e-113;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 935 AAAAAGTATGGCGGAAAAATGACAAATCAGACAGCCGCGGCCCAAGAAATTCGACAAATT 994
Db 1 AAAAAGTATGGCGGAAAAATGACAAATCAGACAGCCGCGGCCCAAGAAATTCGACAAATT 60

QY 995 ATTGGAAATTTGAAGCTGAAATGCGCAGAAATTTGGGATACGAAATGTCAGATTGAC 1054
Db 61 ATTGGAAATTTGAAGCTGAAATGCGCAGAAATTTGGGATACGAAATGTCAGATTGAC 120

QY 1055 AGAACGACATCTGACATTTCTAGATTGTGCGAGGAAATCTCTTTGTTTATAAAGTCAC 1114
Db 121 AGAACGACATCTGACATTTCTAGATTGTGCGAGGAAATCTCTTTGTTTATAAAGTCAC 180

QY 1115 TGGTAAATCGGACAGAGGAGAAATGCAAAAAGTACGATCTACATTTGTTCAAAATCTA 1174
Db 181 TGGTAAATCGGACAGAGGAGAAATGCAAAAAGTACGATCTACATTTGTTCAAAATCTA 240

QY 1175 TGAGGAAACAAAAGTTCAATTGAGTTTCCCACTACCACTAGTCAAAAGTTAAAAGTGG 1234
Db 241 TGAGGAAACAAAAGTTCAATTGAGTTTCCCACTACCACTAGTCAAAAGTTAAAAGTGG 300

QY 1235 AGCAAAAGATACGCTGTACCAATGGAACATCTTGAAGTTTCATGAGAGCCCAAGATA 1294
Db 301 AGCAAAAGATACGCTGTACCAATGGAACATCTTGAAGTTTCATGAGAGCCCAAGATA 360

QY 1295 CAAGATCGAATGTATCTCGTGATGCAAGCAAGTTTCTAAAGCGAGCTACACGAAACC 1354
Db 361 CAAGATCGAATGTATCTCGTGATGCAAGCAAGTTTCTAAAGCGAGCTACACGAAACC 420

QY 1355 TCAGACTCAAGAAATACCCATAAATGCTGAAAGATTTGGATTTCTCTCTCTGAAGA 1414
Db 421 TCAGACTCAAGAAATACCCATAAATGCTGAAAGATTTGGATTTCTCTCTCTGAAGA 480

QY 1415 GCTAAATTTGTTGAAGATTTGGATTATGCTCCAACTTCAGATGATC 1463
Db 481 GCTAAATTTGTTGAAGATTTGGATTATGCTCCAACTTCAGATGATC 529

RESULT 15
D27223/c
LOCUS CELK002A7R Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
DEFINITION clone yk2a7 3', mRNA sequence.
D27223
D27223.1 GI:521294
EST.
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 449)
REFERENCE Kohara,Y., Mitsuiki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
AUTHORS Tabara,H.
TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1994)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. .449
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk2a7"
/sex="hermaphrodite, male"
/tissue_type="whole animal"

/dev stage="varied"
/clone_lib="Yuji Kohara unpublished cDNA"

ORIGIN

Query Match 12.9%; Score 416.4; DB 9; Length 449;
Best Local Similarity 99.0%; Pred. No. 2.1e-87;
Matches 417; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2763 TTTTCTTGGCATCTCATCATGGTCTCTTGGTACATCTCGTCCAGGACATTACACTGTT 2822
Db 421 TTTTCTTGGCATCTCATCATGGTCTCTTGGTACATCTCGTCCAGGACATTACACTGTT 362

QY 2823 ATGTATGACGATAAAGGAATGAGCAAGATGAAGTCTATAAATGACCTACGACTTGT 2882
Db 361 ATGTATGACGATAAAGGAATGAGCAAGATGAAGTCTATAAATGACCTACGACTTGT 302

QY 2883 TTTTCTCTCTGTAGATGTGCAAAACCCATCTCGTTGCCCTGTTCCGGTTTCATTATGCTCAT 2942
Db 301 TTTTCTCTCTGTAGATGTGCAAAACCCATCTCGTTGCCCTGTTCCGGTTTCATTATGCTCAT 242

QY 2943 TTATCATGTGAAAGGCGAAGAGCTTTATCGAACTTACAGGAACATTACATCGGTGAC 3002
Db 241 TTATCATGTGAAAGGCGAAGAGCTTTATCGAACTTACAGGAACATTACATCGGTGAC 182

QY 3003 TATGCACAGCCACGAGCTCGACACGAAATGGAACATTTTCTCCAACTAAACGTGAAGTAC 3062
Db 181 TATGCACAGCCACGAGCTCGACACGAAATGGAACATTTTCTCCAACTAAACGTGAAGTAC 122

QY 3063 CCTGGAATGTGCTTCGCATAACATTTTGCAAAAGTGTGCGCCGTTTCAATCAAAATTTTTC 3122
Db 121 CCTGGAATGTGCTTCGCATAACATTTTGCAAAAGTGTGCGCCGTTTCAATCAAAATTTTTC 62

QY 3123 AATTGTAGATATTGTACTCTTTTAAAGCCGGTTTCAAAATTTCAATCCATGAC 3182
Db 61 AATTGTAGATATTGTACTCTTTTAAAGCCGGTTTCAAAATTTCAATCCATGAC 2

QY 3183 T 3183
Db 1 T 1

Search completed: July 6, 2006, 01:42:43
Job time : 14574 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 21:25:13 ; Search time 17487 Seconds
(without alignments)
11800.671 Million cell updates/sec

Title: US-10-645-746-2
Perfect score: 3227
Sequence: 1 cagcccaaaagtgtgaac.....tttaaaaaaaaaaaaaaa 3227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vi.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_om.*
15: gb_da.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3227	100.0	3229	13	AF180730 Caenorhab
2	2602	80.6	39339	13	CEK08H10
3	760.2	23.6	41648	13	CBRG18K16
4	129.6	4.0	30591	13	CEZK218
5	74.8	2.3	44042	4	SPCC736
6	69	2.1	4583	13	DQ178241
7	69	2.1	29820	13	CEB48F7
8	68.8	2.1	25850	13	CEZK757
9	68.2	2.1	33706	13	AF016682
10	64.8	2.0	4339	11	AJ719770
11	63.4	2.0	3090	4	AK118258
12	62.4	1.9	734	2	BD214386
13	62.4	1.9	734	2	AR767980
14	62.4	1.9	3040	5	BC025769
15	62.4	1.9	3050	2	BD158050
16	62.4	1.9	3050	2	AX879833
17	62.4	1.9	3050	5	AK022827
18	62	1.9	7218	2	I66494

19	61.8	1.9	1966	4	AK221864
20	61.4	1.9	771	2	AR551688
21	61	1.9	1196	4	AK060907
22	61	1.9	1196	4	AK104539
23	61	1.9	1885	4	AK119248
24	61	1.9	3178	4	AK065274
25	61	1.9	3385	4	AK122078
26	60.2	1.9	2897	4	AK122078
27	59.2	1.8	38463	13	CEK01A6
28	58	1.8	3904	5	HSB803237
29	57.6	1.8	2780	5	AB081473
30	56.4	1.7	2144	5	AB168706
31	56.4	1.7	3105	5	AK000040
32	56.4	1.7	5990	5	AB046787
33	55.4	1.7	1165	4	AB236789
34	54.8	1.7	2644	6	AB081474
35	54.8	1.7	3080	6	AK129395
36	54.8	1.7	3700	6	BC096023
37	54.6	1.7	879	4	AY072819
38	54.6	1.7	1211	4	AK101842
39	54.6	1.7	1407	4	AY323483
40	54.2	1.7	1141	2	AR579680
41	54.2	1.7	1141	2	AX083744
42	54.2	1.7	82312	12	AP007836
43	53.8	1.7	4129	11	BC098982
44	53.4	1.7	1522	13	AY701231
45	53	1.6	87618	4	AC007654

ALIGNMENTS

RESULT 1
AF180730
LOCUS
DEFINITION
Caenorhabditis elegans RNA interference promoting factor RDE-1
(rde-1) mRNA, complete cds.
ACCESSION
AF180730
VERSION
AF180730.1
KEYWORDS
GI:6272677
ORGANISM
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 3229)
Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A.,
Timmons, L., Fire, A. and Mello, C.C.
The rde-1 gene, RNA interference, and transposon silencing in C.
elegans
JOURNAL
Cell 99 (2), 123-132 (1999)
PUBMED
10535731
REFERENCE
2 (bases 1 to 3229)
Tabara, H., Sarkissian, M., Kelly, W.G., Grishok, A., Timmons, L.,
Fire, A. and Mello, C.C.
Direct Submission
JOURNAL
Submitted (25-AUG-1999) Medical School, Program in Molecular
Medicine, University of Massachusetts, 373 Plantation Street,
Worcester, MA 01605, USA
FEATURES
Location/Qualifiers
1..3229
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/chromosome="V"
/clone="yk296b10"
1..3229
/gene="rde-1"
21..3083
/genes="rde-1"
/notes="similar to Caenorhabditis elegans K08H10.7"
/codon_start=1
/product="RNA interference promoting factor RDE-1"
/protein_id="AAF06159.1"

/db_xref="GI.6272678"
/translation="MSGNPPELEKGFYRHSLSLDPENMKWLARTGCRKDFKFEYKQLLLV
NWPFKSAYIDREYSEYVKMTKVLNRPQKGFPPKTEIETIDPRAKLFYMQHLRHEKQ
QYTFPLEDVFBSKDTVSEANRYSITSMVSEYKVAKSDSEKKEKQELKLYTMI
LTVKQHLNFSESNPEKDBEARSYKFLNMTQKVPAYVNEEIKVQKPAKVFVD
NNSILRVPSFHPNRPQESLVAPELAEWFGIYIGIELEDFGSEVPNFAIDKLVN
AKWMSLDYLLILLVDFQSDNDVDKLYKLMAGNKTQRAARPTQLLENLKLKCA
EVDNMSRITREHLTLDLDCENSILYKVKTCGRGNRAAKYDTTFLPYEENKFK
EFPHLPVAVKSGAYAPVMEHLEVHKPQYKRNKLDVMQDKLRATKPHDYKE
NTLMLKELFDSSEELNVPFVRLGSLKMQIPCEKVLKVEKSNVNEQITQMTVPYR
GFOBKOLNVPELKKCCAVVNEVETAGPCEBENDVAFYFTLEIGCKCFQIRIGANE
NRGAQSYMTDAKTNEVAYKNTLNTGIRGTEIATAEAKNFERLPDKEQKGMFLIIT
SKQOLNAYGVFKYCDHTIGVANOHLTSETVTYKALASIRHEKSGRYOLAKINAK
LGGINGLDSWEIATLSPEKERRKMTPTMTVDVTPITSYSGIDYSIAAPVASIN
PGGTINYMNIUTOECPGERVAHGRKPTDILEAKFVKLIREFAENNDRAPAHIVIN
YRQGVDSSEMLRVSHDLRLSLKSEYKQFMSRSDKGTVPKFSSTIVIQKRNHLLRME
EKQPKVYNKLTPTAETDVAVAQWEEEDMKESKETGPNSSGTTVDKLIIVSKYKFL
FHLASHGLVLTGSRGHYTMVDYDKGMSQDVEYKMTYGLAFGLASCRKPTSLSPVPVHY
AHLSCAKELXETPKYEHYIGDQAPRTEBHEVLOTNNYKPGMSPA

ORIGIN

Query Match	100.0%;	Score 3227;	DB 13;	Length 3229;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3227;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CAGCCACAAAGTGATGAACATGCTCCTCGAATTTTCCCGAATTTGGAAAGAGATTTTATC	60	
Db	1	CAGCCACAAAGTGATGAACATGCTCCTCGAATTTTCCCGAATTTGGAAAGAGATTTTATC	60	
Qy	61	GTCAATTTCTCGATCCGAGATGAATGGCTTTCGAGGCCCACTGGTAAATCGACGGCA	120	
Db	61	GTCAATTTCTCGATCCGAGATGAATGGCTTTCGAGGCCCACTGGTAAATCGACGGCA	120	
Qy	121	AATTTCTATGAGAAAGTACTCTTTTGGTAAATTTGGTTTCAAGTTTCCAGCAAAATTT	180	
Db	121	AATTTCTATGAGAAAGTACTCTTTTGGTAAATTTGGTTTCAAGTTTCCAGCAAAATTT	180	
Qy	181	ACGATCGGGAATACTACGAGTATGAAGTGAAATCACAAGGAAGATTTTGAATAGAAAC	240	
Db	181	ACGATCGGGAATACTACGAGTATGAAGTGAAATCACAAGGAAGATTTTGAATAGAAAC	240	
Qy	241	CAGGAAAACTTTTCCAAAAAGACAGAAAATTTCCAAATTTCCCGATCGTCAAAACTCTTCT	300	
Db	241	CAGGAAAACTTTTCCAAAAAGACAGAAAATTTCCAAATTTCCCGATCGTCAAAACTCTTCT	300	
Qy	301	GGCAACATCTTCGGCATGAGAAGAGACAGATTTTATTTCTCGAAGACTATGTTTTG	360	
Db	301	GGCAACATCTTCGGCATGAGAAGAGACAGATTTTATTTCTCGAAGACTATGTTTTG	360	
Qy	361	ATGAAAAGGACACTGTTTATAGTGTGTGCGACTGAACACTGTCACATCAAAAATGCTCG	420	
Db	361	ATGAAAAGGACACTGTTTATAGTGTGTGCGACTGAACACTGTCACATCAAAAATGCTCG	420	
Qy	421	TTTCGGGAAAAAGTAGTAAAAAGGATTCGGAGAAAAAGATGAAAAGGATTTGGAGAAA	480	
Db	421	TTTCGGGAAAAAGTAGTAAAAAGGATTCGGAGAAAAAGATGAAAAGGATTTGGAGAAA	480	
Qy	481	AAATCTTATACAAATGATCTTACTCTATCGTAAAAAAATTTTCCACCTGAACCTTTAGTCGAG	540	
Db	481	AAATCTTATACAAATGATCTTACTCTATCGTAAAAAAATTTTCCACCTGAACCTTTAGTCGAG	540	
Qy	541	AAAAATCCGGAAAAAGACGAAGAGGAAATCCGGAGTTTACAAAATTCCTGAAGAATGTTATCGA	600	
Db	541	AAAAATCCGGAAAAAGACGAAGAGGAAATCCGGAGTTTACAAAATTCCTGAAGAATGTTATCGA	600	
Qy	601	CCCAGAAAAGTTTCGCTACGCGCCTTTTGTGGAACGAGGAGATTAAAGTACAAATTCGCGAAAA	660	
Db	601	CCCAGAAAAGTTTCGCTACGCGCCTTTTGTGGAACGAGGAGATTAAAGTACAAATTCGCGAAAA	660	
Qy	661	ATTTTGTGTACGATTAATTAATTCAAATCTGCGAGTTCCTGAAATCGTTTTACGATCCAAACA	720	
Db	661	ATTTTGTGTACGATTAATTAATTCAAATCTGCGAGTTCCTGAAATCGTTTTACGATCCAAACA	720	

Db 1801 |||||TCGCTAGATTGAAATAGCCCAACAGAGGCAAGATATGTTGAAGCTTTCCCGGATA 1860
Qy 1861 AAGAACAAAAGTCTTAATGTTTCATTATCAATTTCCAAACGACAACTGAATGCTTTACGGTT 1920
Db 1861 AAGAACAAAAGTCTTAATGTTTCATTATCAATTTCCAAACGACAACTGAATGCTTTACGGTT 1920
Qy 1921 TTGTGAACACATATTGCGATCAGACCACTCGGTGAGCTTAATCAGCATATTACTTCTGAAA 1980
Db 1921 TTGTGAACACATATTGCGATCAGACCACTCGGTGAGCTTAATCAGCATATTACTTCTGAAA 1980
Qy 1981 CAGTCACAAAAGCTTTGGCATCACTAAGGCGACGAGAAAGGATCAAAACGAAATTTTCTATC 2040
Db 1981 CAGTCACAAAAGCTTTGGCATCACTAAGGCGACGAGAAAGGATCAAAACGAAATTTTCTATC 2040
Qy 2041 AAATTGCAATTGAAATCAACGCGAAATTAAGAGGATTAATACCGAGGCTTGACTGTCAG 2100
Db 2041 AAATTGCAATTGAAATCAACGCGAAATTAAGAGGATTAATACCGAGGCTTGACTGTCAG 2100
Qy 2101 AAATTGCGAAGATATCACCGAGAAAGAAAGAAAGACGGAACCAATGCCATTAACTATGT 2160
Db 2101 AAATTGCGAAGATATCACCGAGAAAGAAAGAAAGACGGAACCAATGCCATTAACTATGT 2160
Qy 2161 ATGTTGGAATTGATGAATCACTCAACCTCTACAGTGGAAATTTGATTTCTATAGCGG 2220
Db 2161 ATGTTGGAATTGATGAATCACTCAACCTCTACAGTGGAAATTTGATTTCTATAGCGG 2220
Qy 2221 CTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCTCGAAATATGATTTGACTCAAG 2280
Db 2221 CTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCTCGAAATATGATTTGACTCAAG 2280
Qy 2281 AAGAAATGCTCCGGTGAGCGTGCAGTGCATGACGCGGAAAGAAACAGATATTTTGG 2340
Db 2281 AAGAAATGCTCCGGTGAGCGTGCAGTGCATGACGCGGAAAGAAACAGATATTTTGG 2340
Qy 2341 AAGCAAAAGTTGCGTGAATGCTCAGAGAAATTCGAGAAACCAACGACAAATCGAGCACAG 2400
Db 2341 AAGCAAAAGTTGCGTGAATGCTCAGAGAAATTCGAGAAACCAACGACAAATCGAGCACAG 2400
Qy 2401 CGCATATTTGATGCTATCGAGACGAGTATAGCGATTCGAGAGATGCTACGTGTTAGTCATG 2460
Db 2401 CGCATATTTGATGCTATCGAGACGAGTATAGCGATTCGAGAGATGCTACGTGTTAGTCATG 2460
Qy 2461 ATGAGCTTCATCTTTAAAAGCGAGATGAAAACAAATTCATGTCGGAACGGGATGAGAAAG 2520
Db 2461 ATGAGCTTCATCTTTAAAAGCGAGATGAAAACAAATTCATGTCGGAACGGGATGAGAAAG 2520
Qy 2521 ATCCAGAGCGAAGTACACGCTTCATTTGATTCAGAAACACACAAATACACGATTTGCTTC 2580
Db 2521 ATCCAGAGCGAAGTACACGCTTCATTTGATTCAGAAACACACAAATACACGATTTGCTTC 2580
Qy 2581 GAAGAAATGGAAGAAAGATAAGCCAGTGTCTAATAAGATCTTACTCTGCTGAAACAGATG 2640
Db 2581 GAAGAAATGGAAGAAAGATAAGCCAGTGTCTAATAAGATCTTACTCTGCTGAAACAGATG 2640
Qy 2641 TCCTGTTGCTGCTTTAAACATGAGGAGAGATGAAAGAAAGAAAGAAACATGGA 2700
Db 2641 TCCTGTTGCTGCTTTAAACATGAGGAGAGATGAAAGAAAGAAAGAAACATGGA 2700
Qy 2701 TTGTGAACCATCATCCGGAACCACTGTGATTAACCTTATCGTTTCGAAATACAAATTCG 2760
Db 2701 TTGTGAACCATCATCCGGAACCACTGTGATTAACCTTATCGTTTCGAAATACAAATTCG 2760
Qy 2761 ATTTTTTCTTGGCATCTCATCATGTTGCTCTTGGTACATCTCGTCCAGGACATTAACACTG 2820
Db 2761 ATTTTTTCTTGGCATCTCATCATGTTGCTCTTGGTACATCTCGTCCAGGACATTAACACTG 2820
Qy 2821 TTATGATGACGATAAGGAATGAGCCAGATGAGATGCTATAAATGACCTACGGAATTG 2880
Db 2821 TTATGATGACGATAAGGAATGAGCCAGATGAGATGCTATAAATGACCTACGGAATTG 2880
Qy 2881 CTTTCTCTCTGCTAGATGTCGAAACCCCATCTCGTTGCTGCTGTTCCGGTTTCATTATGCTC 2940
|||||

Db 2881 CTTTCTCTCTGCTAGATGTCGAAACCCCATCTCGTTGCCCTGTTCCGGTTTCATTATGCTC 2940
Qy 2941 ATTATCATGTGAAAAAGCGAAAGAGCTTTATCGAACTTACAAAGGAAACATTACATCGGTG 3000
Db 2941 ATTATCATGTGAAAAAGCGAAAGAGCTTTATCGAACTTACAAAGGAAACATTACATCGGTG 3000
Qy 3001 ACTATGACAGCCACGAGCTCGACACGAAATGGAACATTTTCTCCTCAAACTAACGCTGAAGT 3060
Db 3001 ACTATGACAGCCACGAGCTCGACACGAAATGGAACATTTTCTCCTCAAACTAACGCTGAAGT 3060
Qy 3061 ACCCTGGAATGCTGTCGCATTAACATTTTGCAAAAGTGTCCCGTTTCAATCAAAATTTT 3120
Db 3061 ACCCTGGAATGCTGTCGCATTAACATTTTGCAAAAGTGTCCCGTTTCAATCAAAATTTT 3120
Qy 3121 TCAATTGTAGATATTGTACTTACTTTTTTTTAAAGCCCGGTTTCAAAAATTCATTCCCATG 3180
Db 3121 TCAATTGTAGATATTGTACTTACTTTTTTTTAAAGCCCGGTTTCAAAAATTCATTCCCATG 3180
Qy 3181 ACTAACGTTTTTCATAAATTAATCTTGAATTTTAAAAAATTTAAAAAATTTAAAAA 3227
Db 3181 ACTAACGTTTTTCATAAATTAATCTTGAATTTTAAAAAATTTAAAAAATTTAAAAA 3227

RESULT 2
CEK08H10/c

LOCUS CEK08H10 39339 bp DNA linear INV 09-AUG-2005
DEFINITION Caenorhabditis elegans Cosmid K08H10, complete sequence.
ACCESSION Z83113

VERSION Z83113.1 GI:3217648
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans

REFERENCE 1 (bases 1 to 39339)
C. elegans Sequencing Consortium
Genome sequence of the nematode C. elegans: a platform for
investigating biology
Science 282 (5396), 2012-2018 (1998)

JOURNAL PUBLISHED 9851916
REFERENCE 2 (bases 1 to 39339)
Gardner, A.E.
Direct Submission
Submitted (27-NOV-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
worm@sanger.ac.uk

COMMENT On Jun 13, 1998 this sequence version replaced gi:1695070.
Coding sequences below are predicted from computer analysis, using
predictions from Genefinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
For a graphical representation of this sequence and its analysis
see: [http://www.wormbase.org/perl/ace/elegans/seq/sequence?](http://www.wormbase.org/perl/ace/elegans/seq/sequence?name=K08H10;class=Sequence)
name=K08H10;class=Sequence

This sequence is the entire insert of clone K08H10. The true left
end of clone K0688 is at 32990 in this sequence. The true right end
of clone C2H6 is at 10790 in this sequence. The start of this
sequence (1..104) overlaps with the end of sequence Z81042.
The end of this sequence (39233..39339) overlaps with the start of
sequence Z73975.

FEATURES
source Location/Qualifiers
1..39339 /organism="Caenorhabditis elegans"

QY 1713 AATGAAACAGAGGCGCAATCTATTATGTAGCGCGACGAAATGAATATGCC--- 1769
 Db 2586 AATGAAACAGAGGCGCAATCTATTATGTAGCGCGACGAAATGAATATGCCGTA 2527
 QY 1770 -----TTCTACAAATAATGTACAC 1788
 Db 2526 AGTTTCAGAAATTTGAAGTTTTTAATATCATATTACAGTTCTACAAATAATGTACAC 2467
 QY 1789 TAAATACCGGAATCGGTAGATTGGAATAGCCGCAACAGAGCGAAGAAATATGTTGAAC 1848
 Db 2466 TAAATACCGGAATCGGTAGATTGGAATAGCCGCAACAGAGCGAAGAAATATGTTGAAC 2407
 QY 1849 GTCTTCCCGATTAAGAACAAAGCTTTAATGTTGATTTATCTTCCAAACGCAACTGA 1908
 Db 2406 GTCTTCCCGATTAAGAACAAAGCTTTAATGTTGATTTATCTTCCAAACGCAACTGA 2347
 QY 1909 ATGCTTACCGTTTTGTGAAACATTATTGCGATCACACCATCGGTGATGCTAATCAGCAT 1968
 Db 2346 ATGCTTACCGTTTTGTGAAACATTATTGCGATCACACCATCGGTGATGCTAATCAGCAT 2287
 QY 1969 TTACTTCTGAAACAGTCAAAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2028
 Db 2286 TTACTTCTGAAACAGTCAAAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2227
 QY 2029 GAATTTCTCATCAATTTGCAATGGAATCAACGCGAATTTAGGAGTATTACCGAGGC 2088
 Db 2226 GAATTTCTCATCAATTTGCAATGGAATCAACGCGAATTTAGGAGTATTACCGAGGC 2167
 QY 2089 TTGACTGGTCAGAAATTCAGAAATATCACAGAGAAAGAAAGACGGAAGAAACAAATGC 2148
 Db 2166 TTGACTGGTCAGAAATTCAGAAATATCACAGAGAAAGAAAGACGGAAGAAACAAATGC 2107
 QY 2149 CATTAATATGATGTTGGAATTTGATGTAATCTCAATCCACCTCTACAGTGAATTTGATT 2208
 Db 2106 CATTAATATGATGTTGGAATTTGATGTAATCTCAATCCACCTCTACAGTGAATTTGATT 2047
 QY 2209 ATCTATAGCGCTGATGAGGATATCAATCCAGTGGAACTATCTATCGAAATATGA 2268
 Db 2046 ATCTATAGCGCTGATGAGGATATCAATCCAGTGGAACTATCTATCGAAATATGA 1987
 QY 2269 TTGTGACTCAAGAAATGTCGTCGGGAGCGGTGAGCTGCTCATGACGCGGAAGAA 2328
 Db 1986 TTGTGACTCAAGAAATGTCGTCGGGAGCGGTGAGCTGCTCATGACGCGGAAGAA 1927
 QY 2329 CAGATATTTTGAAGCAAAAGTTGCGTGAATTTGCTCAGAAATTCGCGAA----- 2378
 Db 1926 CAGATATTTTGAAGCAAAAGTTGCGTGAATTTGCTCAGAAATTCGCGAAAGTGAGTTGTC 1867
 QY 2379 -----AACACGAC 2387
 Db 1866 TTGAGTATTTAAAGATCTCGGATTTTTTAATTTTTTTTGTAACTTTTCAGAAACACGAC 1807
 QY 2388 AATCGAGCACCGCATATTGTAGTCTATCGAGCGGAGTTAGCGATTCGGAGATGCTA 2447
 Db 1806 AATCGAGCACCGCATATTGTAGTCTATCGAGCGGAGTTAGCGATTCGGAGATGCTA 1747
 QY 2448 CGTGTGATCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAAATTCATGTCGAA 2507
 Db 1746 CGTGTGATCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAAATTCATGTCGAA 1687
 QY 2508 CGGGATGGAAGATCCAGCGCGAGTACAGTTGATTTGATTTCCAGAAAGACACAT 2567
 Db 1686 CGGGATGGAAGATCCAGCGCGAGTACAGTTGATTTGATTTCCAGAAAGACACAT 1627
 QY 2568 ACACGATGCTTCGAAGATGGAAGAAAGATAGCCAGTGTCAATAAGATCTTACTCCT 2627
 Db 1626 ACACGATGCTTCGAAGATGGAAGAAAGATAGCCAGTGTCAATAAGATCTTACTCCT 1567
 QY 2628 GCTGAAACAGATGTCGCTGTTGCTGTTGTTAAACAAATGGAGGAGGATATGAAGAAAGC 2687
 Db 1566 GCTGAAACAGATGTCGCTGTTGCTGTTTAAACAAATGGAGGAGGATATGAAGAAAGC 1507

QY 2688 AAGAAACTGGAATTTGTAACCCATCATCCGAAACAACTGTGGATAAACTTATCGTTTCG 2747
 Db 1506 AAGAAACTGGAATTTGTAACCCATCATCCGAAACAACTGTGGATAAACTTATCGTTTCG 1447
 QY 2748 AATACAAATTCGATTTTTTTTCTTGGCATCTCATATGTTGCTCTTGGTATCATCTCGTCCA 2807
 Db 1446 AATACAAATTCGATTTTTTTTCTTGGCATCTCATATGTTGCTCTTGGTATCATCTCGTCCA 1387
 QY 2808 GGACATTACACTGTTATGATGACCAATAAGGAATGAGCCAGATGAAGTCTAT----- 2861
 Db 1386 GGACATTACACTGTTATGATGACCAATAAGGAATGAGCCAGATGAAGTCTATGTAAGC 1327
 QY 2862 ----- 2861
 Db 1326 GTTTGAATAGCAGTAGCGATTTTAGGATTTTGTAAATCCGCATATAGTTATATATAAAA 1267
 QY 2862 -----AAATGACCTACGACTTGTCTCTCTGTAGATGTCGAAACCCAT 2911
 Db 1266 AATGTTTCAGAAATAGACCTACGGACTTGTCTCTCTGTAGATGTCGAAACCCAT 1207
 QY 2912 CTCGTTGCTGTTCCGTTTATTATGCTCATTTATCATGTAAGAAAGCGAAGAGCTTTA 2971
 Db 1206 CTCGTTGCTGTTCCGTTTATTATGCTCATTTATCATGTAAGAAAGCGAAGAGCTTTA 1147
 QY 2972 TCGAACTTTACAGGAACATTATCATCGGTGACTATGACACGACCGACTCGACACGAAAT 3031
 Db 1146 TCGAACTTTACAGGAACATTATCATCGGTGACTATGACACGACCGACTCGACACGAAAT 1087
 QY 3032 GGAACTTTTCCAAACTTAACGTGAAGTACCTCGAATGTCGTTGCGATATACATTTTGC 3091
 Db 1086 GGAACTTTTCCAAACTTAACGTGAAGTACCTCGAATGTCGTTGCGATATACATTTTGC 1027
 QY 3092 AAAAGTGTGCGCGTTCAATCAATTTTCAATTTTCAATTTGATAGATTTGACTTTTTTTT 3151
 Db 1026 AAAAGTGTGCGCGTTCAATCAATTTTCAATTTTCAATTTGATAGATTTGACTTTTTTTT 967
 QY 3152 AAAGCCCGTTTCAAAATTTCAATTCATGACTAAGCTTTTCAATAAATTAATTTGAATTTA 3211
 Db 966 AAAGCCCGTTTCAAAATTTCAATTCATGACTAAGCTTTTCAATAAATTAATTTGAATTTA 907

RESULT 3
 CBRG18K16
 LOCUS 41648 bp DNA linear INV 04-NOV-2000
 DEFINITION Caenorhabditis briggsae cosmid G18K16, complete sequence.
 AC084520
 AC084520.1 GI:11094970
 VERSION HTG.
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE 1 (bases 1 to 41648)
 AUTHORS Washington University Genome Sequencing Center.
 TITLE The C. briggsae Genome Sequencing Project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 41648)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA
 e-mail: jsp1eth@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we provide a small overlap between
 neighboring submissions.
 Location/Qualifiers

[illegible]

gene	<p>/translation="MLSVSIKPIFVGHIIYSITYETWIGLLSIDRFLSSRSABPNRFR TORSYFYLLFIFFYAKELGFIWIVSIVSDSKNKLQVIFYFYTSYIFLQIILFI GMIFOLTKSSQDQLTRQTKIIGATKGLFVLFGIVTGFVVVSTIFASIDPILV SVMLTEIMCSPSPGTGIN" join(9736, 9849, 9903, .10158,10207, .10586) locus_tag="ZK218.5" join(9736, .9849,9903, .10158,10207, .10586) locus_tag="ZK218.5" /standard_name="ZK218.5" /note="contains similarity to Pfam domain PF01549 (SHTK domain)" /codon_start=1 /product="Hypothetical protein ZK218.5" /protein_id="CAB04986.1" /db_xref="GI:3881561" /db_xref="InterPro:IPR003582"</p>
CDS	<p>/translation="MVASIVVELALASCVQAQPAVGTALNCTVFDATVQDFVYTPG AVNCNLYADTKTCETIYNPTEDAPVPAKCFALAAATDVEPLDKVTKAA SCPCTGVCSSPDYKCPNVQFPLNCATILPSCRDQWRITIAQDPSACGFCNQ GCVDAIVCANDISICNAVGQDFVNLNQCQTKNRCSTTTVGGIIGCSTYNRSSNA CAAAVNGFCQNFVTPQRKSYCATTCRIC" complement(join(12056, .12300,12359, .12731,12774, .12976, 13024, .13231)) /gene="nhr-248" locus_tag="ZK218.6" complement(join(12056, .12300,12359, .12731,12774, .12976, 13024, .13231)) /gene="nhr-248" locus_tag="ZK218.6" /standard_name="ZK218.6" /note="contains similarity to Pfam domains PF00104 (ligand-binding domain of nuclear hormone receptors), PF00105 (Zinc finger, C4 type (two domains))" /codon_start=1 /product="Hypothetical protein ZK218.6" /protein_id="CAB04989.1" /db_xref="GI:3881564" /db_xref="GOA:O46003" /db_xref="HSSP:LA6Y" /db_xref="InterPro:IPR000536" /db_xref="InterPro:IPR001628"</p>
gene	<p>/translation="MKSIIQCPILSISSPSVLFEACVCGPPSSRRFGALACFCRN VFFRMLNLKNCISNCPCEITGKTRYACQCFKKCKLKGMDPKALKHRDKIGPR SPKMSKNKPQSAEIDPFCILLIDQKSKHQVYNQYRLAKESDNLMLRIGFM NATBWCNGFRPRYSVKRILISBEFGIASGLIDQAFKTATATONEWLFNNCLR PDAC7VHNDPLNTLKRELSQPFILGDSPECVILKTLMLTPSPFQSKLQSTQKS RDKTSELKCLAKFPLSGVERGEIILLIGSIRITVKFEHNHTRKRSOLINPNHFDI PVTSPFLMA" join(14387, .14509,14555, .14813,14863, .15260) locus_tag="ZK218.7" join(14387, .14509,14555, .14813,14863, .15260) locus_tag="ZK218.7" /standard_name="ZK218.7" /note="contains similarity to Pfam domain PF01549 (SHTK domain)" /codon_start=1 /product="Hypothetical protein ZK218.7" /protein_id="CAB04987.1" /db_xref="GI:3881562" /db_xref="InterPro:IPR003582"</p>
CDS	<p>/translation="MASSIVIVLALASCVNAAAPTIGTELNCTVFLATLTPDADFVY TPGATCNVSYSDSTCAVYITDEBPVGNTPILPKCFAGTDGTVGADVPDLVR AAVSTCPKTCGLCCQSDADYKCPNVQFPLNCATILPSCRDQWRVIAQDPCPSAGF CNOQGGVDATVICSNDISICNAVGQDFVNLNQCQTKGRCSTTTAAGSVATASGTC TYNADNSRSCAAWAANGFCNTPISTAQRSSCATTCRIC" complement(join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" /standard_name="ZK218.11" /note="contains similarity to Pfam domain PF01549 (SHTK domain)" /codon_start=1</p>
gene	<p>/translation="MASSIVIVLALASCVNAAAPTIGTELNCTVFLATLTPDADFVY TPGATCNVSYSDSTCAVYITDEBPVGNTPILPKCFAGTDGTVGADVPDLVR AAVSTCPKTCGLCCQSDADYKCPNVQFPLNCATILPSCRDQWRVIAQDPCPSAGF CNOQGGVDATVICSNDISICNAVGQDFVNLNQCQTKGRCSTTTAAGSVATASGTC TYNADNSRSCAAWAANGFCNTPISTAQRSSCATTCRIC" complement(join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" /standard_name="ZK218.11" /note="contains similarity to Pfam domain PF01549 (SHTK domain)" /codon_start=1</p>
CDS	<p>/translation="MASSIVIVLALASCVNAAAPTIGTELNCTVFLATLTPDADFVY TPGATCNVSYSDSTCAVYITDEBPVGNTPILPKCFAGTDGTVGADVPDLVR AAVSTCPKTCGLCCQSDADYKCPNVQFPLNCATILPSCRDQWRVIAQDPCPSAGF CNOQGGVDATVICSNDISICNAVGQDFVNLNQCQTKGRCSTTTAAGSVATASGTC TYNADNSRSCAAWAANGFCNTPISTAQRSSCATTCRIC" complement(join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" join(16134, .16531,16581, .16830,16897, .17061)) locus_tag="ZK218.11" /standard_name="ZK218.11" /note="contains similarity to Pfam domain PF01549 (SHTK domain)" /codon_start=1</p>

```

/product="Hypothetical protein ZK218.11"
/protein_id="CAB04990.1"
/db_xref="GI:3881565"
/db_xref="InterPro:IPR003582"
/translation="MSSIKNGKLARFFFTPSFAMAFIISLAVALSQVAVPTIGV
ELMCTVFGDFTFYIPATNCVNTSYSDTKTDITYIATPAVQVFGASTARPKCYAAA
ATDPNPDLVSTVATCTPKTCGFCQSDYNCNPNRPLNCDTILSCXKQQRWRI
IADPCASCFNCGQGVDAI PCPNDLSICNNVGMDFNLNCKQKTCGRCPSTTPTR
TGGVTPGGTCTTYNADSSRTCAAWANGFCIKNEFYSITQKASCATTCLIC"
gene join(24630..24740,24787..25125,25263..25333,25577..25613)
locus_tag="ZK218.8"
join(24630..24740,24787..25125,25263..25333,25577..25613)
CDS join(24630..24740,24787..25125,25263..25333,25577..25613)
locus_tag="ZK218.8"
/standard_name="ZK218.8"
/notes="contains similarity to Pfam domain PF02171 (Pwi
domain)"
/codon_start=1
/product="Hypothetical protein ZK218.8"
/protein_id="CAB04988.1"
/db_xref="GI:3881563"
/db_xref="InterPro:IPR003165"
/translation="MDKSHVSYRPPTRRLGLYFPVSATGCVLVVDTIKIHDILKTVN
FEGLHYVNAAGTLEYWASEYNYTIKIKNHFVYSLAPLSKKNPKVKQSETPRFINPFS
GTVDKLVGSGYKDFYLNHSHAVLGTSRPAHYTVMYDDMGMSQDEVYKMTDALFLS
ARCKPISLPAPVHFYVYKLFQCPW"
gene complement(29297..29394)
snRNA complement(29297..29394)
locus_tag="ZK218.13"
complement(29297..29394)
locus_tag="ZK218.13"
/product="Hypothetical RNA transcript ZK218.13"

Query Match 4.08; Score 129.6; DB 13; Length 30591;
Best Local Similarity 73.78; Pred No. 1.1e-25;
Matches 165; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 2860 ATAAATGACCTACGGACTTGCTTTCTCTCTGTAGATGTCGAAAACCCATCTCGTTGC 2919
Db 25261 AGAAATGACTGAGGCTCTTGCTCTCTCGGCAGATGCGAAAACCCATCTCCCTTC 25320

Qy 2920 CTGTTCCGGTTCATTATGCTCATTTATCATGTGNAAGGCGAAGAGCTTTATCGAAGCT 2979
Db 25321 CTGCCCCGGTTCAGTAGCTCTTTGTCGCGAAGAGGCTTAAAGACCAATATGAGCGCT 25380

Qy 2980 ACAAGGAACATTACATCGGTGACTATGCACAGCCAGGACTCGACACGAATGGAACATT 3039
Db 25381 ACATGAGCACTTTATTGACGCGACACACCACTAATTCGACGGAATGGAAGAT 25440

Qy 3040 TTCTCCAAACTAACGTGAAGTACCCTCGAATGCTGTTCCGAT 3083
Db 25441 TTCTCCAAAGCAATTCGATTATCCCGAATTTGTTGTGTA 25484

RESULT 5
SPCC736
LOCUS S.pombe chromosome III cosmid c736.
DEFINITION
ACCESSION AL023705
VERSION AL023705.1 GI:3169070
KEYWORDS
SOURCE
ORGANISM Schizosaccharomyces pombe (fission yeast)
Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
1 (bases 1 to 44042)
Wood,V., Gwilliam,R., Rajandream,M.A., Lyne,M., Lyne,R.,
Stewart,A., Sgouros,J., Peat,N., Hayles,J., Baker,S., Basham,D.,
Bowman,S., Brooks,K., Brown,D., Brown,S., Chillingworth,T.,
Churcher,C., Collins,M., Connor,R., Cronin,A., Davis,P.,
Fellwell,T., Fraser,A., Gentles,S., Goble,A., Hamlin,N., Harris,D.,
Hidalgo,J., Hodgson,G., Holtroyd,S., Hornsby,T., Howarth,S.,
Huckle,E.J., Hunt,S., Jagels,K., James,K., Jones,L., Jones,M.,
Leather,S., McDonald,S., McLean,J., Mooney,P., Moule,S.,

```

Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neill, S., Pearson, D., Quail, M.A., Rabinowitz, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Meltjens, I., Vanstreels, E., Rieger, M., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritze, C., Holzer, E., Moestl, D., Hilbert, H., Borzým, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Furnelle, B., Goffeau, A., Cadieu, B., Dreano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garçon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Foréburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G. and Nurse, P.

The genome sequence of *Schizosaccharomyces pombe*
 Nature 415 (6874), 871-880 (2002)

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Details of *S. pombe* sequencing at the Sanger Institute are available on the World Wide Web.
 (URL, <http://www.genedb.org/genedb/pombe/index.jsp>) (URL, http://www.sanger.ac.uk/Projects/S_pombe/)
 CDS are numbered using the following system eg SPAC5H10.01c. SP (S. pombe), A (Chromosome 1), c5H10 (cosmid name), .01 (first CDS), c (complementary strand). However, clones may have been reorientated since the original submission, therefore the complementary strand notation may be invalid for strand inference. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once.

FEATURES

source

1. .44042

/organism="Schizosaccharomyces pombe"

/mol_type="genomic DNA"

/strain="972h"

/db_xref="taxon:4896"

/chromosome="III"

/map="III"

/clone="cosmid c736"

1. .1158

/gene="SPCC553.01c"

/note="synonym: SPCC736.01c"

1. .1158

/gene="SPCC553.01c"

/codon_start=1

/protein_id="CAA19265.1"

/db_xref="GI:3169071"

/translation="DQROQSSVAEPSPTSKRLRLRLPLLRKPTGRQNFINTVTPSKNLYRSVKTAELASLQNVTTTNEIPNSPDFLDAASEDEHPIPLSDSSLLTFLQKNKAFQVPRVPSFGSPPLVKPASSSFGSHSLGTSLRNSLNGSSASSITFLQNNVNPVSPSTVLEHAKHPSLAKSHGSLTGLLIYPKYANGSKRGDGCATSIISGHSLSLPKNSRNVSPSPFNPVPSFTVSSDQVQKKGNVPNAIETDSSPSDTISSPTVRANVLSKQKPKPKLTPIDIFSAGKSKSPILPTQESCSLPIDIEFS

complement(1. .181)

/note="nominal overlap with cosmid SPCC553, EM:AL023704 S.

pombe chromosome 3"

1655. .2515

/gene="SPCC736.02"

1655. .2515

/gene="SPCC736.02"

/note="hypothetical protein; sequence orphan"

/codon_start=1

misc_feature

complement(1. .181)

/note="nominal overlap with cosmid SPCC553, EM:AL023704 S.

pombe chromosome 3"

1655. .2515

/gene="SPCC736.02"

1655. .2515

/gene="SPCC736.02"

/note="hypothetical protein; sequence orphan"

/codon_start=1

/protein_id="CAA19266.1"

/db_xref="GI:3169072"

/db_xref="UniProtKB/TrEMBL:O74951"

/translation="MGKKRKTKINDYNEKITRKVKKNLTDGKARALTVFBSLPLEV
 LRLIFLLNNSLAVTSRSLRRLSRNNTPIFMPIDFTLSMVPKSIILSIQGLLRR
 YFTLQILTKIDELVTSFGKSPNEDIKDGRVIFHAGHEGFIKRLILFLNEDFIAEL
 EHRNFLKASSLRNGFLALKORNPVTRQGVKIVTERLISIPDDEVEFCFTWPEYT
 BEQNSVELLDIVFGPWNILIEKRFKSFIDSVLTRLVDIAISKVNTEIWNLYIEKAI
 POLPNLLKLF"

complement(2836. .4125)

/gene="SPCC736.03c"

complement(2836. .4125)

/gene="SPCC736.03c"

/note="phenylalanine-tRNA synthetase; localization
 mitochondrion (predicted); similar to S. cerevisiae MSF1"

/codon_start=1

/protein_id="CAA19267.1"

/db_xref="GI:3169073"

/db_xref="GOA:O74952"

/db_xref="InterPro:IPR002319"

/db_xref="InterPro:IPR004530"

/db_xref="InterPro:IPR005121"

/db_xref="InterPro:IPR006195"

/db_xref="UniProtKB/TrEMBL:O74952"

/translation="MPPHRVSHLSKLFRRFPNKRKQVTDSSNVPFIHISKIGRNL
 FQKEGHPICSLRQLLEQFOKPEMNNVOKESPIVSVETNFDLSLGFPPKTHVSRKSDTY
 YNNKTCILRTHSAHQPEFSLAKEGKNGFLITADVYRDEVDSSHVPIPHOMEG
 ALVNRNDTAKMNDLKVSLKSMVDDTETALETHNDQDIYSLETKEFVSQHL
 KDTLTTVHDLVSLAPSISGAEQVRYRTYDSFPFKPSQLEIDMKKWLIEGCG
 VQDRLKAGLNNYIGWAFGIGLERLAMILYIPDIRLFWLSDERFSKQFIPNKLST
 FKFSKYPACFKDIAFWINNDPNDFEIIIRDVQDMVESVNLIDQYTKSGKTSLC
 YRVNYSRSLRNEIDKLOELRRNVANSVELR"

complement(3184. .3402)

/gene="SPCC736.03c"

/note="Match to PF01409 tRNA-synt_2d, tRNA synthetases
 class II (F) Score 80.50"

complement(3208. .3237)

/gene="SPCC736.03c"

/note="PS00339 Aminoacyl-transfer RNA synthetases

class-II signature 2"

complement(3646. .3978)

/gene="SPCC736.03c"

/note="Match to PF01409 tRNA-synt_2d, tRNA synthetases

class II (F) Score 29.47"

complement(3649. .3702)

/gene="SPCC736.03c"

/note="PS00179 Aminoacyl-transfer RNA synthetases

class-II signature 1"

complement(5197. .6324)

/gene="gmal2"

/note="synonym: SPCC736.04c"

complement(5197. .6324)

/gene="gmal2"

/EC_number="2.4.1.-"

/note="galactosyltransferase;
 alpha-1,2-galactosyltransferase (PMID 95003210);
 GNA12/MNN10 family; localization golgi (PMID 95003210);
 involved in galactosylation of cell surface glycoproteins
 (PMID 9839953); transfers galactose to both N-linked and
 O-linked oligosaccharides (PMID 9839953); predicted
 N-terminal signal sequence; similar to S. pombe gmh1 and
 gmh2 and gmh3 and SPBC1289.13C and SPAC5H10.13C and
 SPAC637.06 and SPBC82.17 (paralogues); similar to S.
 cerevisiae MNN10 and MNN11"

/codon_start=1

/protein_id="CAA19268.1"

/db_xref="GI:3169074"

/db_xref="GOA:O09174"

/db_xref="InterPro:IPR008630"

/db_xref="UniProtKB/Swiss-Prot:O09174"

/translation="MRFAPYLISAVVITIIIGGAWTSAMDTKLOTWKKEIIDQHTS
 TWTPVSSVTSTQDTLRVTITSEVSVTATLETFTATFTVTSVVALATTDPHPDPS
 KIVILMGSNFQDANSPLHPPAQSIKNNREYABRGYKFEFLDADAYASRVTHLMP"


```

RESULT 7
CEP48F7
LOCUS      29820 bp      DNA      linear      INV 09-AUG-2005
DEFINITION Caenorhabditis elegans Cosmid F48F7, complete sequence.
ACCESSION  Z69661
VERSION     Z69661.1 GI:1200275
KEYWORDS   HTG; Syntaxin like.
SOURCE     Caenorhabditis elegans
ORGANISM   Caenorhabditis elegans
            Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
            1 (bases 1 to 29820)

REFERENCE
  1. C. elegans Sequencing Consortium
    Genome sequence of the nematode C. elegans: a platform for
    investigating biology
    Science 282 (5396), 2012-2018 (1998)
    9851916
  2. (bases 1 to 29820)
    Coles, L.
    Direct Submission
    Submitted (24-FEB-1996) Nematode Sequencing Project, Sanger
    Institute, Hinxton, Cambridge CB10 1SA, England and Department of
    Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
    wormesanger.ac.uk

COMMENT
  Coding sequences below are predicted from computer analysis, using
  predictions from GeneFinder (P. Green, U. Washington), and other
  available information.
  Current sequence finishing criteria for the C. elegans genome
  sequencing consortium are that all bases are either sequenced
  unambiguously on both strands, or on a single strand with both a
  dye primer and dye terminator reaction, from distinct subclones.
  Exceptions are indicated by an explicit note.
  IMPORTANT: This sequence is NOT necessarily the entire insert of
  the specified clone. It may be shorter because we only sequence
  overlapping sections once, or longer because we arrange for a small
  overlap between neighbouring submissions.
  For a graphical representation of this sequence and its analysis
  see: - http://www.wormbase.org/perl/ace/elegans/seq/sequence?name=F48F7;class=Sequence
  IMPORTANT: This sequence is not the entire insert of clone F48F7.
  It may be shorter because we only sequence overlapping sections
  once, or longer because we arrange for a small overlap between
  neighbouring submissions.
  The true left end of clone F48F7 is at 1 in this sequence. The true
  right end of clone F48F7 is at 12044 in
  sequence Z78539.
  The true left end of clone C31E10 is at 29717 in this sequence. The
  true right end of clone K02B9 is at 5215 in this sequence. The
  start of this sequence (1..104) overlaps with the end of sequence
  Z69663.
  The end of this sequence (29717..29820) overlaps with the start of
  sequence Z78539.

FEATURES
  Source
    1..29820
    Location/Qualifiers
      /organism="Caenorhabditis elegans"
      /mol_type="genomic DNA"
      /strain="Bristol N2"
      /db_xref="taxon:6239"
      /chromosome="X"
      /clone="F48F7"
      join(Z69663.1:30546..30755,112..284,431..601,655..2117,
      2173..3125,3178..3216)
      /genes="alg-1"
      /locus tag="F48F7.1"
      join(Z69663.1:30546..30755,112..284,431..601,655..2117,
      2173..3125,3178..3216)
      /genes="alg-1"
      /locus tag="F48F7.1"
      /standard name="F48F7.1"
      /note="C. elegans ALG-1 protein; contains similarity to
      Pfam domains PF02170 (PAZ domain), PF02171 (Piwi domain)"

gene
  /codon_start=1
  /product="Hypothetical protein F48F7.1"
  /protein_id="CAA93496.2"
  /db_xref="GI:22265872"
  /db_xref="InterPro:IPR003100"
  /db_xref="InterPro:IPR003165"
  /db_xref="UniProtKB/TrEMBL:Q20578"
  /translation="MSGPQYLFGVWNSTIQQQPSATSSFLSPGSPISSTSSQVVP
  TSGATQPPPPSAAGAAATALQNDLEIFNSPQTQPTFSDVPOQAGSLAPGVPIGN
  TSVSIGEPANTLGGGLPGGAPGOLPGNQSGIOFQCRPRRNHVEGRSILIRANHFV
  RIPGRTIOHYQVDVTPDKCPRVNRREIISCLISAPSKYFTNIRVPYQGRNMVTRPL
  PIGRMDPVTLPDGSNAVERQPSVSLKWQVSLSDLEDAWGRVQRQVPEAQQAMD
  VIIRHLPSLKTPTVGRSFFSPFPVNSAGWAGSCFPQASGAVAGAHAGSAGYHAESKL
  GGGREVMFGHQSRVSPQKMLNIDVSATAFYRSMVEVIEFIAEVLPLVQALAEERA
  LSDAQRVKFTKEIRGLKIEITHGQMRKRVNNTVRRPAQTQTFPIQLETGTQIECT
  VAKYFYDKYRQLKPHLPCLOVGOEKHTYLPPEVCNIPVGGRCIKKLTPVQSTMI
  KATARSAPREREISNLVRKAESADPFAHEFGITINPAMTEVGRVLSAPKLLYGGR
  TRATALPNQGVDMRGKQFHTGIDVRWALACPAQQQHVXENDLRMTNQLQRLSNDRA
  GMPIVGNPCFKYAVGVEQPEPMFKYLQNYSGIQLVVVLPKGTPIYAEVRYRGDVT
  LGIATQVQAKNAIRTPPTLSNLCLKNYKLGAGVNSILLPNVRPIFNEVPIFGCD
  ITHPPAGDSRKSIAAVGVGSMDAHPSRYAATVRQOHRQBIISDLTYMRELLVQFGR
  NTRFKPARIVVYRDGVSEGOFFENVLOYELRAIIRACMMLERGVOPGTTTAVQKRHT
  RLFAVDKQOVGKAYNIPGTTVDVIGITHEPDPYLCSHAGIQGTSRPSHYHVMDD
  NNTLDELQLTVMCHTVYRCRTRSPAPATYAHLVAFRARTHLVDRHDSGEGSQ
  PSTGSDTTLSSNARAVQVHPDANNVYFA"
  join(8676..8702,9134..9220,9295..9396,9446..9545,
  10008..10219,10265..10417,10465..10581,10842..10943)
  /genes="syn-2"
  /locus tag="F48F7.2"
  join(8676..8702,9134..9220,9295..9396,9446..9545,
  10008..10219,10265..10417,10465..10581,10842..10943)
  /genes="syn-2"
  /locus tag="F48F7.2"
  /standard name="F48F7.2"
  /note="C. elegans SYN-2 protein; contains similarity to
  Pfam domains PF05739 (), PF00804 (Syntaxin)"
  /codon_start=1
  /product="Hypothetical protein F48F7.2"
  /protein_id="CAA93492.4"
  /db_xref="GI:54110860"
  /db_xref="GOA:Q20574"
  /db_xref="HSSP:JUTH"
  /db_xref="InterPro:IPR0000727"
  /db_xref="InterPro:IPR006011"
  /db_xref="InterPro:IPR006012"
  /db_xref="UniProtKB/Swiss-Prot:Q20574"
  /translation="MEDRLNEFSQSVTDRDFEVELSPARPSPSAEYVDRFEFVRNAL
  ASVRGIEKLRRDQOHVLTADPRDKNILENQIGTIRRTGDLRKLVRQAEDDFLE
  FTQVQSIVTEKMRQQLBLLKDNLNKLINLFNETHDYKRSVRRVRQLQVGRDGL
  TDIEDINRMENSGEQLFFREVNPLSVGQAAYEDVKRGEIKDLNNLTAMLEEIFL
  DLQHLTEAQDEMVTNIDNNVENGLEQVKGQSANVKTAVEYKKSAMRKKICVAAILIT
  LLIIITVAIILAVLSRGNNKK"
  join(12943..12981,13029..13154,13201..13313,13361..13565,
  13887..14059,14104..14209)
  /locus tag="F48F7.3"
  join(12943..12981,13029..13154,13201..13313,13361..13565,
  13887..14059,14104..14209)
  /locus tag="F48F7.3"
  /standard name="F48F7.3"
  /note="contains similarity to Pfam domain PF01762
  (galactosyltransferase)"
  /codon_start=1
  /product="Hypothetical protein F48F7.3"
  /protein_id="CAA93493.1"
  /db_xref="GI:3877296"
  /db_xref="GOA:Q20575"
  /db_xref="InterPro:IPR002659"
  /db_xref="UniProtKB/TrEMBL:Q20575"
  /translation="MORENFGPWLNNNAVMYFIVGTDIDVIDDIDDEIKTYEDIIQVD
  INENYHNTYKAIFWINEINKRCHGPKLLKDDVDHIDMIGIQFLIKRYSVMDDFMA
  CTVISNGVVRNNTSKWYLSKKDYKFPNLTGYCOGMVYFISGNLLPILHGNIEKQPL
  WNDWYVTRSLVGDYKISYISLEQHSISPNTVREYVSSYSLMAIKTRKWTITFAHFRFPQ
  KPFMYRRRKIWANITEINNCSNILLRKARINFVSAY"

```


QY	2896	GATGTGAAACCCATCTCGTTCCTGTTCCGGTTCATTATGCTCATTT	2944
Db	2959	GATGCACACGATCCGTTTCAATTCACGCGCCAGCATATTATGCTCATTT	3007
RESULT 8			
CEZK757/c			
LOCUS	CEZK757	25850 bp	DNA linear INV 09-AUG-2005
DEFINITION	Caenorhabditis elegans Cosmid ZK757, complete sequence.		
ACCESSION	Z29121		
VERSION	Z29121.1	GI:438366	
KEYWORDS	HTG.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;		
	Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.		
REFERENCE	1	(bases 1 to 25850)	
AUTHORS	C. elegans Sequencing Consortium		
CONSTRM	Genome sequence of the nematode C. elegans: a platform for		
TITLE	Investigating biology		
JOURNAL	Science	282 (5396), 2012-2018 (1998)	
PUBMED	9851916		
REFERENCE	2	(bases 1 to 25850)	
AUTHORS	Thomas, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JAN-1994)	Nematode Sequencing Project, Sanger	
COMMENT	Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: worm@sanger.ac.uk		
	Coding sequences below are predicted from computer analysis, using		
	predictions from GeneFinder (P. Green, U. Washington), and other		
	available information.		
	Current sequence finishing criteria for the C. elegans genome		
	sequencing consortium are that all bases are either sequenced		
	unambiguously on both strands, or on a single strand with both a		
	dye primer and dye terminator reaction, from distinct subclones.		
	Exceptions are indicated by an explicit note.		
	IMPORTANT: This sequence is NOT necessarily the entire insert of		
	the specified clone. It may be shorter because we only sequence		
	overlapping sections once, or longer because we arrange for a small		
	overlap between neighbouring submissions.		
	For a graphical representation of this sequence and its analysis		
	see:- http://www.wormbase.org/perl/ace/elegans/seq/sequence?		
	name=ZK757;class=Sequence		
	IMPORTANT: This sequence is not the entire insert of clone ZK757.		
	It may be shorter because we only sequence overlapping sections		
	once, or longer because we arrange for a small overlap between		
	neighbouring submissions.		
	The start of this sequence (1..414) overlaps with the end of		
	sequence Z22180.		
	The end of this sequence (25695..25850) overlaps with the start of		
	sequence Z30215.		
FEATURES	Location/Qualifiers		
source	1..25850		
	/organism="Caenorhabditis elegans"		
	/mol_type="genomic DNA"		
	/strain="Bristol N2"		
	/db_xref="taxon:6239"		
	/chromosome="III"		
	/clones="ZK757"		
gene	join(complement(3000..3053), complement(1928..2009),		
	complement(1666..1787), complement(1415..1623),		
	complement(1285..1362), complement(1038..1128),		
	complement(888..950), complement(684..798),		
	complement(113..239), complement(222180.1:32171..32424),		
	complement(222180.1:31663..31679))		
	/locus_tag="ZK757.4a"		
	join(complement(3000..3053), complement(1928..2009),		
	complement(1666..1787), complement(1415..1623),		
	complement(1285..1362), complement(1038..1128),		
	complement(888..950), complement(684..798),		
CDS			

gene	complement(join(14326..14392,14451..15271,15331..15760,15814..16025,16084..16308,16370..16450,16726..16893,16942..17165,17346..17558,17741..18077))	
	/gene="pqn-39"	
	/locus_tag="F48F7.4"	
	complement(join(14326..14392,14451..15271,15331..15760,15814..16025,16084..16308,16370..16450,16726..16893,16942..17165,17346..17558,17741..18077))	
	/gene="pqn-39"	
	/locus_tag="F48F7.4"	
	/standard_name="F48F7.4"	
	/note="C. elegans PQN-39 protein; contains similarity to Saccharomyces cerevisiae Substrate of the Hub1 ubiquitin-like protein that localizes to the shmoo tip (mating projection); mutants are defective for mating projection formation, thereby implicating Hbtlp in polarized cell morphogenesis; SGD:YDL223C"	
	/codon_start=1	
	/product="Hypothetical protein F48F7.4"	
	/protein_id="CAA93494.1"	
	/db_xref="GI:3877297"	
	/db_xref="UniProtKB/TrEMBL:Q20576"	
	/translation="MLCFSCRSLFQFNFLLLSILFLFTNSHLLITISLSSSSTPWCFLQRRRRASVERNVCLVSWACGRORNMVKGPTTALLIVVISGRGVDDGN EGSRRGIARSLPVIVLCLSLSQPIELQSOYISVFCVQFOFANOILLQSLGLLKH FKLYSTFLFRCSKLLTYQIPACNSSFLVMSAKODLAPATVSHADTGLQLPP QHQTDPNTHGPIIYQSIAPVAHYQGPQVWNPQPSVTHVYQAPDASSQEN KQVSNHYTHYQAQMDGMLRPQQAQYVQQYVTPPQYPAQQTQTFISYA TTFYANDYTPHNMGRGAGRDDYHVAQNSVSGVMQVFPVAPVGNVETQAHQ SANSYANVSSYSTIIPENMKDPESDNI SOYDQRTSOLNGENTDAEDHLSPPTVQ NHPFNVEPRPOONFVEKESQSTTRMSKNSKGRDELTFDERLHKRMQKS RVNENGVGATSPATTVAAPSNVANSTSLTPRSNQCKKETSGSVRNEGSNNQ NNTKRVGNAGHMYNGEHSAGHEQFQQQQQRYVYNGATGERQRPRNDQ SKRVYNNHNSNNRSPYQAQQNQPTQONQHPREGAPENYRQOSYRGNRGG GAYRGYQHQPQHLQNPLOQPLNQYPPAQIMGPQWATMNAFQNAVITGLPPG VORIPLPGFPMGPROVHODIKFQNTWLDATSCPLPAAVMSIDTRFGRGMRGRGV VANYTVGAYRGVGYAAPRYIGRPVETPATPKESSEOSTEVDMDAKKDDERSVQ DRASATESVETPADQGTQEQNKQSDSQNGQEPVAVKXANBEIILLPANTVQSPD AAAREKVEST"	
gene	join(24169..24244,24310..24416,24468..24656,24863..25019,25077..25429)	
	/locus_tag="F48F7.5"	
CDS	join(24169..24244,24310..24416,24468..24656,24863..25019,25077..25429)	
	/locus_tag="F48F7.5"	
	/standard_name="F48F7.5"	
	/note="contains similarity to Pfam domain PF00096 (Zinc finger, C2H2 type)"	
	/codon_start=1	
	/product="Hypothetical protein F48F7.5"	
	/protein_id="CAA93490.1"	
	/db_xref="GI:3877293"	
	/db_xref="GOA:Q20572"	
	/db_xref="InterPro:IPR007087"	
	/db_xref="UniProtKB/TrEMBL:Q20572"	
	/translation="MADIPKRIFDIWQPKLIEFTCGTVMDVYKLOSKCKDEKENTENS FAIHDIUKNAERDVHYQVQHARLLMSQNLHAWFCGASQNEEDLKGLCKEKFEN	
Query Match	2.1%;	Score 69; DB 13; Length 29820;
Best Local Similarity	56.3%;	Pred. No. 3.9e-08;
Matches 129;	Conservative 0;	Mismatches 100; Indels 0; Gaps 0;
QY	2716	CCGGAACAACCTGGATAAATCTATCGTTTCGAAATACAAATTCGATTTTCTTGGCAT 2775
Db	2779	CAGGAACACTGTCGATAGTAGGAATACCCACCAACCGAGTTTGACTTCTACTTGTCT 2838
QY	2776	CTCATCATGTGTCCTTGGTACATCTCGTCCAGGACATACACTGTTATGTATGACGATA 2835
Db	2839	CTCATGCTGTATTCAAGGAACATCTCGTCCATCCCATACCAATGTTCTTGGATGACA 2898
QY	2836	AAGGAATGACCGCAAGATGAAGTCTATAAAATGACCTACGGACTTGTCTTCTCTGCTA 2895
Db	2899	ACAATCTGACAGCGCGAACTTCAACAGCTCACATATCAAAATGTGCCATACCTACGTCA 2958


```

complement (113..239), complement (222180.1:32171..32424),
complement (222180.1:31663..31679))
/locus_tag="ZK757.4a"
/standard_name="ZK757.4a"
/notes="contains similarity to Pfam domain PF01529 (DHHC
zinc finger domain)"
/codon_start=1
/product="Hypothetical protein ZK757.4a"
/protein_id="CAD54174.2"
/db_xref="GI:24817642"
/db_xref="GOA:Q810G4"
/db_xref="InterPro:IPR001594"
/db_xref="UniProtKB/Swiss-Prot:Q810G4"
/translations="MSRCRCKGCGVTFWLRFLPVLVSLATGWIYATYELC
ILSDNPQRILYIFVYALLILFTYLRVTKAWPKQKCIEGASATYEVK
DEQLQFLSDIARDITLLVRGDFHGRFCDCCKCIPDRSHSCMKQKDFD
HCFVNNVNFNGYKYRFLFLAYGFICWIAATLPSFIDFWRHEYDMCKQVDSID
SVLQIKHLHTVLSNGRFLVLLFLSCMFLSLSPFLFVHLVLTAKNRTTVESPA
PMIDGKADKAFNHGIRANYREIFGSHPLWFLPVRSISGDGCKFVNDMTASAAQ
NQPVENGVPNGQSHAPPIAQHHINLYSEQSSASEKEIKSVDDDEITERSQVSTAT
TASPSHDSVTA"
join (complement (2427..2486), complement (1928..2009),
complement (1666..1787), complement (1415..1623),
complement (1285..1362), complement (1038..1128),
complement (888..950), complement (684..798),
complement (113..239), complement (222180.1:32171..32424),
complement (222180.1:31663..31679))
/locus_tag="ZK757.4b"
/locus_tag="ZK757.4b"
join (complement (2427..2486), complement (1928..2009),
complement (1666..1787), complement (1415..1623),
complement (1285..1362), complement (1038..1128),
complement (888..950), complement (684..798),
complement (113..239), complement (222180.1:32171..32424),
complement (222180.1:31663..31679))
/locus_tag="ZK757.4b"
/standard_name="ZK757.4b"
/codon_start=1
/product="Hypothetical protein ZK757.4b"
/protein_id="CAH19096.1"
/db_xref="GI:51587415"
/db_xref="GOA:Q65ZC4"
/db_xref="InterPro:IPR001594"
/db_xref="UniProtKB/TrEMBL:Q65ZC4"
/translations="MCDVFKYKLDLILKYLVRFLPVLVSLATGWIYATYVE
LCILSDNPQRILYIFVYALLILFTYLRVTKAWPKQKCIEGASATYEV
KDERQLQFLSDIARDITLLVRGDFHGRFCDCCKCIPDRSHSCMKQKLVF
DHRCFVNNVNFNGYKYRFLFLAYGFICWIAATLPSFIDFWRHEYDMCKQVDS
IDSVIQRNLHLHTVLSNGRFLVLLFLSCMFLSLSPFLFVHLVLTAKNRTTVESF
RAPMIDGKADKAFNHGIRANYREIFGSHPLWFLPVRSISGDGCKFVNDMTASAA
AGNQPVENGVPNGQSHAPPIAQHHINLYSEQSSASEKEIKSVDDDEITERSQVST
ATTASPSHDSVTA"
complement (join (3966..4183,4237..4368,4674..4764,
4849..4944,5337..5432,5479..5658,5713..5818,6332..6387))
/locus_tag="ZK757.1"
/locus_tag="ZK757.1"
complement (join (3966..4183,4237..4368,4674..4764,
4849..4944,5337..5432,5479..5658,5713..5818,6332..6387))
/locus_tag="ZK757.1"
/standard_name="ZK757.1"
/notes="contains similarity to Pfam domain PF06681 ()"
/codon_start=1
/product="Hypothetical protein ZK757.1"
/protein_id="CAA82387.3"
/db_xref="GI:24817641"
/db_xref="GOA:P34679"
/db_xref="InterPro:IPR010601"
/db_xref="UniProtKB/Swiss-Prot:P34679"
/translations="MIYNNISLNTIYTLTYRDLTFPFSVHGTCVFNFLFCIPFI
CRPLHRTKPTIFFVLGTFVLSLFLFLQFVYVFLWSLPERVTIAVCLVKCIT
SSTTSQAQLPLAVAIYRPIVVRNKKMPSFWVHVSIFFFIVAIINFLPLGSE
TNDQCAVLRFSKAMEARISLTGLNLFAVINAIYFVKYDKRNDVHRRRVOLT
YVMLQSMIPILVSIPLLVGSFDFYFGLFWELLWIKCKKISCFIHRFYPSWLHVS
LVNNITSPSLADNLNFAFTSNNSPRTVLYHFLFELFIHNSNIKFSY"
complement (join (9579..9810,10116..10180,10227..10370),

```

```

10695..10794,11316..11422,11470..11603,12543..12645))
/locus_tag="ZK757.2"
/standard_name="ZK757.2"
/notes="contains similarity to Pfam domain PF00782 (Dual
specificity phosphatase, catalytic domain)"
/codon_start=1
/product="Hypothetical protein ZK757.2"
/protein_id="CAA82388.2"
/db_xref="GI:30424335"
/db_xref="GOA:P34680"
/db_xref="InterPro:IPR000340"
/db_xref="InterPro:IPR000387"
/db_xref="UniProtKB/Swiss-Prot:P34680"
/translations="MPSRRRIINREORQCSQIVRLVYVSLAALSPRVLSRFCVCINLI
PGRFLSAPPHKVVHLPLQNETDITDLSPHWANYKEIEBARKGGRALLCMMGISRS
ATFGIATVMQYERKTLHDSYKAVQLARNIICPNVFPFQQLIDLEQLRGKVSCKIIEP
LPCKQPDVTWQELYDEMISMQDHRSLASCNLSARSTNDTMSLRSLMNVNDTSR
SLASFHLTRPIGASPTLLVPSSSSSSSVRGPIPLQRAHTEPPKEASILPKSALRDKS
KSSEKKKWRLSFHKDVV"
join (complement (230215.2:861..915),
complement (230215.2:201..373), complement (25676..25848),
complement (24734..25627), complement (24473..24689),
complement (23916..24423), complement (23604..23866),
complement (23434..23558), complement (22889..23386),
complement (22717..22841), complement (22579..22670))
/genes="tag-76"
/locus_tag="ZK757.3a"
join (complement (230215.2:861..915),
complement (230215.2:201..364), complement (25676..25848),
complement (24734..25627), complement (24473..24689),
complement (23916..24423), complement (23604..23866),
complement (23434..23558), complement (22889..23386),
complement (22717..22841), complement (22579..22670))
/genes="tag-76"
/locus_tag="ZK757.3a"
/standard_name="ZK757.3a"
/notes="contains similarity to Pfam domains PF02170 (PAZ
domain), PF02171 (Pwi domain)"
/codon_start=1
/product="Hypothetical protein ZK757.3a"
/protein_id="CAA82389.1"
/db_xref="GI:3981773"
/db_xref="InterPro:IPR003100"
/db_xref="InterPro:IPR003165"
/db_xref="UniProtKB/Swiss-Prot:P34681"
/translations="MSRRNATSFVNDNLTSSIGSGSMSPPISTRPASGOASPLTS
NGSLSPVADQGSVSNLSDPSRLSLSELACLNREVVARPGLGTIGRIQIPVK
SNFPAMDLANPKMVTVOHVEIHHPCBCKLDKDEWFIIFWKAUSDHNPINHPKALAY
DGAHQLYTVARLEFPDDQGSVRLDCASLIPKONDRTRCAISIQNVGVLLENQRTRT
NLNDEKRVLTPIQLDIIICQSLTCEPLLNKSNANPYTWKSCYRIPTAAGQALDLGGKRE
MWTGFPSSAHIASNRYRPLNIDVAHTAFYKTRITVILQFMCVDLNLNERTSKPRNRRGP

```

```

Query Match      2.1%; Score 68.8; DB 13; Length 25850;
Best Local Similarity 55.4%; Pred. No. 4.4e-08;
Matches 133; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 2705 GAACCCATCATCGGAACAACCTGTGGATAAATATCGTTTCGAAATACAAATTCGATTT 2764
      |||||
Db 23167 GAACGTTCCACCTGGTACACAGTTCACACTGGAATCGTCTCTCCCGAAGGATTGATTT 23108
      |||||

QY 2765 TTTCCTGGCATCTCATCGTGTGCTTGGTACATCTCGTCCAGGACATTACACTGTAT 2824
      |||||

```

Db 23107 CTATCTGTGCTCTCACTACGGAGTACAGGAACTTCTCGTCTCGAATATCATGTTCT 23048

Qy 2825 GTATGACGATAAGGAATGAGCAAGATGAAGTCTATATAAATGACCTACGGACTTGCTTT 2884

Db 23047 TCTGGATGAATGCAAGTCTCACTGCTGATGAATCCAAAGCATCACCTACGGAATGTGTCA 22988

Qy 2885 TCTCTCTGCTAGATGCGAAACCCATCTCGTGTGCTGTTCCGGTTCATTATGCTCATTT 2944

Db 22987 TACATATGGTGGTGTACTCGTTCGTTCCATTCCTCAACTCCAGTTTATTATGCTGATTT 22928

RESULT 9

AF016682 33706 bp DNA linear INV 22-SEP-2004

LOCUS Caenorhabditis elegans cosmid T07D3, complete sequence.

DEFINITION AF016682

ACCESSION AF016682

VERSION AF016682.2 GI:49035142

KEYWORDS HTG.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 33706)

AUTHORS WormBase Consortium

CONSTRM Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium Science 282 (5396), 2012-2018 (1998)

JOURNAL 9851916

PUBMED 2 (bases 1 to 33706)

AUTHORS Lamar, B., Wamsley, P. and Twyman, B.

TITLE The sequence of C. elegans cosmid T07D3

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (31-JUL-1997) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 4 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (25-JUL-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 5 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (21-SEP-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 6 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 7 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (23-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 8 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (29-OCT-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE 9 (bases 1 to 33706)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

10 (bases 1 to 33706)

Waterston, R.

Direct Submission

Submitted (10-FEB-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

11 (bases 1 to 33706)

Wilson, R.

Direct Submission

Submitted (22-JUN-2004) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

12 (bases 1 to 33706)

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL

COMMENT

Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1HQ, England
email: submissions@watson.wustl.edu and jesse@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this clone sequence and its analysis see:
<http://www.wormbase.org/db/seq/sequence?name=T07D3;class=Sequence>

NEIGHBORING CLONE INFORMATION

The 5' clone is Y46B2A, 200 bp overlap; the 3' clone is R07C3, 200 bp overlap. Actual start of this clone is at base position 1 of T07D3; actual end is at 2312 of R07C3.

NOTES:

Coding sequences below are the result of integration and manual review of the following data : computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES

source

1..33706
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"

```

/gene
/chromosome="II"
/clonage="T07D3"
750..6014
/genes="alg-2"
/locus tag="T07D3.7"
join(750..804,1698..1800,2398..2571,2641..2981,3273..4334,
4566..5524,5976..6014)
/genes="alg-2"
/locus tag="T07D3.7"
/standard name="T07D3.7a"
/notes="contains similarity to Pfam domains PF02170 (PAZ
domain), PF02171 (Piwi domain); coded for by the following
C. elegans cDNAs: OSTF122A8.1, yk81c3.5,
yk109d4.5, yk160h5.5, yk175g11.5, yk185g2.5, yk199g3.5,
yk334f12.5, yk433a5.5, yk464f7.5, yk471e4.5, yk504h5.5,
yk515b9.5, yk519f3.5, yk564d8.5, yk592d7.5, yk609g12.5,
yk635g10.5, yk651b2.5, yk1044a12.3, yk1044a12.5,
yk1063d06.3, yk1063d06.5, yk1290a10.3, yk1290a10.5,
CEESW71R, yk1689c07.3, yk67f3.3, yk81a6.3, yk81c3.3,
yk109d4.3, yk160h5.3, yk334f12.3, yk433a5.3, yk515b9.3,
yk564d8.3, yk609g12.3, yk635g10.3, yk651b2.3, yk679a4.3"
/codon start=1
/product="Argonaute (plant)-like gene protein 2, isoform
a"
/protein_id="AA866197.2"
/db_xref="GI:24418217"
/db_xref="WormBase:T07D3.7a"
/translations="MFPLPVHNGPRLGKLSIFPMPGDSLTSFMPDGGAEFTSSSQL
GSAHGAIGTKPDAGVQFCPPVNPNGHVEGRSILLRANHPAVRIPGGSVOHYOIVDPP
CDKPRVNRVIEICLLSSFSKYFTNRPVVDGKRMVTRPLPIGTEPMNFVETLPGD
SAVERFSVETMWIGVCLSLADDDAMEGRVQVPEAVQSIDVILRHLPSLKYTPVGR
SFTPTPGVMPGMQVQESKLGGRVFGFQSHVSPQSMNMLINDSATAFRAMP
VIFVAEVLPEVQALAEARLSDAQRVFTKIRGLKIEITHCGAVRRKRYVCNVR
RPAQTFPLQLEGTGTICTEYAKYFDFKYRIOLKYPHLPCLQVGQEQKHLYLPPEVC
DIVPGORCLKKITDVOTSTMIKATARSAPEREREICLVSKAELSDAPPAHERGITIN
PAMTEYKGRVLSAPKLLYGRHRTATLNPQGVMDRGKQFHTGMVETWJAJACFAQ
SHVENDLMPFTQLORISTDAGNPILGTFMFCYASGVQEVPEPKYLKQTSALQL
IVVLPKGTPIYAEVRKGVGTGLGATQCVQAKNAIRTPQLTSLNCLKNKVLGVQV
SILLNPRRIEVEPIFLGCDITHPAAGTRKPSIAAVGSDMAHPSRYAATRVQV
HQEITITDLTYVRELLQVYRNTFRKPARIVVYRDGVSEGLQFNVLQVELRAIREAC
VLMESGQPGITETIAVQKQHRLFAADKADQVGKAPNIPGGTVDVGTHTPEPFF
LCSHAGIQTSRPSHVLWDDNDLTADLEQLTYOMCHTYVCTSVTSVPAPAYAH
LVAFRARYHLVDRDHSRGESQPSGTSSEDTLSSMAKAVQHPDSNMYFA"
join(1700..1800,2398..2571,2641..2981,3273..4334,
4566..5524,5976..6014)
/genes="alg-2"
/locus tag="T07D3.7"
/standard name="T07D3.7b"
/notes="contains similarity to Pfam domains PF02170 (PAZ
domain), PF02171 (Piwi domain); coded for by the following
C. elegans cDNAs: OSTF122A8.1, OSTRI22A8.1, yk67f3.5,
yk81c3.5, yk109d4.5, yk160h5.5, yk175g11.5, yk185g2.5,
yk199g3.5, yk334f12.5, yk433a5.5, yk464f7.5, yk471e4.5,
yk504h5.5, yk515b9.5, yk519f3.5, yk532d3.5, yk564d8.5,
yk592d7.5, yk609g12.5, yk635g10.5, yk651b2.5, yk1044a12.3,
yk1044a12.5, yk1063d06.3, yk1063d06.5, yk1290a10.3,
yk1290a10.5, CEESW71R, yk1689c07.3"
/codon start=1
/product="Argonaute (plant)-like gene protein 2, isoform
b"
/protein_id="AA038604.1"
/db_xref="GI:28275097"
/db_xref="WormBase:T07D3.7b"
/translations="MPGDSLTSSFPDGGAEFTSSSQLGSAHGAIGTKPDAGVQFQ
CPVRPNHGVGRSILLRANHPAVRIPGGSVOHYOIVDPPDKPRVNRVIEICLLISF
SKYFTNIRVYDGRKNMVTREPLPICTEPMNFVETLPGDSAVEKFSVTKWIGQVCL
SALDDAMEGRVQVPEAVQSIDVILRHLPSLKYTPVGRSFTTPPGVMPGMQVQES
2.1%; Score 68.2; DB 13; Length 33706;
Query Match
Best Local Similarity 55.2%; Pred. No. 6.5e-08;
Matches 133; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
2704 TGAACCATCATCCGGAACACTGTGTGATAAATCTATCGTTTCGAATAACAATTCGATT 2763

```

```

Db 5160 TCACATCCACAGGAACAACACTGTAGATGTGGAATCACGCATCCAAOOGGAATTCGACT 5219
Qy 2764 TTTTCTTGGCATCTCATCATGCTGTCTTGTGTACATCTCGTCCAGGACATTACACTGTGA 2823
Db 5220 TCTTCTCTGCTCCACGCTGTATCCAGGAACATCCGACCATCTCCTACCTACCATGTGC 5279
Qy 2824 TGTATGACGATAAGGAATGAGCAAGATGAAGTCTATATAAATGACCTACGAGCTTGT 2883
Db 5280 TCTGGGATGACAATGATCTGACGCCGACGAGCTTCAACAGTTGACCTACCAAAATGTGC 5339
Qy 2884 TTCTCTCTGATAGTGTGCAAAACCATCTCGTTGTCTCTGCTGTTCGGTTCATTATCTCT 2943
Db 5340 ATACCTATGTAGATGACGAGGTCTGTGTGATCCCGCGCGCGGTACTATGCATC 5399
Qy 2944 T 2944
Db 5400 T 5400

RESULT 10
AJ719770
LOCUS AJ719770 4339 bp mRNA linear VRT 12-JAN-2005
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 6f12.
ACCESSION AJ719770
VERSION AJ719770.1 GI:53129948
KEYWORDS ORF1.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1
Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J.,
Fiedler, P., Kuter, S., Blagodatski, A., Kostovska, D., Kofer, M.,
Flachy, J., Carrini, P., Hayashizaki, Y., and Buerstedde, J.M.,
Full-length cDNAs from chicken bursa lymphocytes to facilitate
gene function analysis
Genome Biol. 6 (1), R6 (2005)
15642098
2 (bases 1 to 4339)
Caldwell, R.B.
Direct Submission
Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuherberg, GERMANY
1, D-85764 Neuherberg, GERMANY
FEATURES
Location/Qualifiers
source
1..4339
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="6f12"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikenl"
/dev_stage="2 weeks old"
1..4339
/locus tag="RCJMB04_6f12"
<1..261
/locus_tag="RCJMB04_6f12"
262..2844
/locus_tag="RCJMB04_6f12"
/notes="ORF1"
/codon start=1
/product="hypothetical protein"
/protein_id="CAC31429.1"
/db_xref="GI:53129949"
/translational="MEISSAGPVGAQPLLMVRRPGVGTMGKPIKLANCFQVEIPKI
DYVLXVDIKPKCKPRVNRVVDVSMVQHPKVTIFGDRPVYDVKSLYTNPLPVAT
TGVDDLTLPFGEGKDRPKVSIKFKVSRVSWLLHLVLTGRTILPEPLPKPDIETNPV
HAVDVVIRLPSNMKYTPVGRSFFSAPEGYDHPUGGREGVWFVGHQSVRPANWKNMLNI

```


TITLE Novel human genes and gene expression products ii
JOURNAL Patent: JP 202519000-A 2528 02-JUL-2002;
COMMENT CHIRON CORP, HISEQ INC

OS Homo sapiens (human)
PN JP 2002519000-A/2528
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580
PR 28-JAN-1998 US 60/072910, 24-FEB-1998 US 60/075954 PR.
31-MAR-1998 US 60/080114, 03-APR-1998 US 60/080515 PR
03-APR-1998 US 60/080666, 21-OCT-1998 US 60/105234 PR
28-OCT-1998 US 60/105877
PI LOUIS T WILLIAMS, JAIME ESCOBEDO, MICHAEL A INNIS, PABLO PI
DOMINGUEZ GARCIA.
PI JULIE SUDDUTH KLINGER, CHRISTOPH REINHARD, KLAUSE GIESE, FILIPPO
PI RANDAZZO,
PI GIULIA C KENNEDY, DAVID POT, ALTAIF KASSAM, GEORGE LAMSON, RADOJE
PI DRMANAC,
PI RADOMIR CRKVENJAKOV, MARK DICKSON, SNEZANA DRMANAC, IVAN LABAT,
PI DENA LESHKOWITZ, DAVID KITA, VERONICA GARCIA, LEE WILLIAM JONES,
PI BIRJIT STACHE CRAIN
PC C12N15/09, C12N15/09, C07K14/47, C07K14/82, C07K16/18, C12N1/15, PC
C12N1/19,
PC C12N1/21, C12N5/10, C12Q1/68, C12N15/00, C12N5/00, C12N15/00 CC n
FH Key = A, T, C or G
FT misc feature Location/Qualifiers
source 1. 734
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.9%; Score 62.4; DB 2; Length 734;
Best Local Similarity 51.8%; Pred. No. 3.7e-06;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCATCATCGGAACAACACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757
|||||
DB 209 GAATGGCAATATCCAGCTGGAACAACAGTTGTATACAGACATTACACACCCATATGAT 268
|||||
QY 2758 TCGATTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 2877
|||||
DB 269 TCGATTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 328
|||||
QY 2818 CTGTTATGTATGACGATAAAGGATGACCAAGATGAAGTCTATAAATGACCTACCGAC 2877
|||||
DB 329 ATGTTTATGGATGATGAATCTCTTTACTGCAGATGAACCTTCAGCTGCTAACTACCGAC 388
|||||
QY 2878 TTGCTTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 2937
|||||
DB 389 TCTGCCACACTTACGTACGCTGTACACGATCTGTTCTATACCTGCACCGCGTATTATG 448
|||||
QY 2938 CTCATTTATCATGTGAAACCGAAGAGCTT 2969
|||||
DB 449 CTCACCTGGTAGCAATTTAGAGCCAGATATCAT 480
|||||

RESULT 13
AR767980 734 bp DNA linear PAT 08-DEC-2005
LOCUS Sequence 2528 from patent US 6964868.
DEFINITION AR767980
ACCESSION AR767980.1 GI:83343259
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Unclassified.
1 (bases 1 to 734)
Williams, L. T., Escobedo, J., Innis, M. A., Garcia, P. D.,
Sudduth-Klinger, J., Reinhard, C., Giese, K., Randazzo, F.,
Kennedy, G. C., Pot, D., Kassam, A., Lamson, G., Drmanac, R.,
Crkvenjakov, R., Dickson, M., Drmanac, S., Labat, I., Leshkowitz, D.,

TITLE Kita, D., Garcia, V., Jones, L. W. and Stache-Crain, B.
JOURNAL Human genes and gene expression products ii
COMMENT Patent: US 6964868-A 2528 15-NOV-2005;
Nuvelo, Inc.; Sunnyvale, CA

FEATURES
source 1. 734
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 1.9%; Score 62.4; DB 2; Length 734;
Best Local Similarity 51.8%; Pred. No. 3.7e-06;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCATCATCGGAACAACACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757
|||||
DB 209 GAATGGCAATATCCAGCTGGAACAACAGTTGTATACAGACATTACACACCCATATGAT 268
|||||
QY 2758 TCGATTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 2817
|||||
DB 269 TCGATTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 328
|||||
QY 2818 CTGTTATGTATGACGATAAAGGATGACCAAGATGAAGTCTATAAATGACCTACCGAC 2877
|||||
DB 329 ATGTTTATGGATGATGAATCTCTTTACTGCAGATGAACCTTCAGCTGCTAACTACCGAC 388
|||||
QY 2878 TTGCTTTTCTCTGTAGATGTGGAACCAAGATGAAGTCTATAAATGACCTACCGAC 2937
|||||
DB 389 TCTGCCACACTTACGTACGCTGTACACGATCTGTTCTATACCTGCACCGCGTATTATG 448
|||||
QY 2938 CTCATTTATCATGTGAAACCGAAGAGCTT 2969
|||||
DB 449 CTCACCTGGTAGCAATTTAGAGCCAGATATCAT 480
|||||

RESULT 14
BC025769 3040 bp mRNA linear PRI 20-OCT-2004
LOCUS Homo sapiens eukaryotic translation initiation factor 2C, 3, mRNA
DEFINITION (CDNA clone IMAGE:5210832).
ACCESSION BC025769
VERSION BC025769.1 GI:22213082
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 3040)
Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,
Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D.,
Altschul, S. P., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K.,
Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,
Scheet, T. E., Brownstein, M. J., Ussid, T. B., Casavant, T. L.,
Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J.,
Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J.,
McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,
Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W.,
Villalón, D. K., Munz, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y.,
Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D.,
Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M.,
Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smal, D. E.,
Schnerch, A., Schein, J. E., Jones, S. J. and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3040)
Director MGC Project.

TITLE
JOURNAL
 Direct Submission
 Submitted (06-MAR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
REMARK
COMMENT
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgebps@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastripop, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgouon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 49 Row: j Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 29294646.

FEATURES

Location/Qualifiers
 1..3040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5210832"
 /tissue_type="Blood, adult leukocytes"
 /clone_lib="NIH_MGC_118"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.9%; Score 62.4; DB 5; Length 3040;
 Best Local Similarity 51.8%; Pred. No. 3.5e-06;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757
 DB 2091 GAAGTGGCAATATCCAGCTGGAAACAAAGTTGATACAGACATTACACACCCATATGAGT 2150
 QY 2758 TCGATTTTCTTGGCATCTCATCTGTTGGTGTCTTGGTACATCTCGTCCAGGACATTACA 2817
 DB 2151 TCGATTTTTCCTCTGTAGCCATGCTGGAATACAGGTTACAGGTACCACTCGTCTTCACACTATC 2210
 QY 2818 CTGTTATGTATGACGATAAAGGAATCAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
 DB 2211 ATGTTTTATGGGATGATACTGCTTTTACTGCAGATGAATCTGCTGCTAACTTACACG 2270
 QY 2878 TTGCTTTTCTCTCTGCTAGATGTGGAAGCCATCTCGTTGCTGTTCGGGTTCAATTATG 2937
 DB 2271 TCTGCCACACTTACGTACGCTGTACAGCATCTGTTCTATACCTTCGACCAGCGTATTATG 2330
 QY 2938 CTCATTATCATGTGGAAGGCGAAAGAGCTT 2969
 DB 2331 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2362

RESULT 15

BD158050
 LOCUS 3050 bp DNA linear PAT 17-JAN-2003
 DEFINITION
 Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION
 BD158050
 VERSION
 BD158050.1 GI:27863808

KEYWORDS
 JP 2002191363-A/12893.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
REFERENCE
 1 (bases 1 to 3050)
AUTHORS
 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE
 Primer for synthesizing full-length cDNA and use thereof
JOURNAL
 Patent: JP 2002191363-A 12893 09-JUL-2002;
COMMENT
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/12893
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
 source
 Location/Qualifiers
 (101)..(2680).
 FT CDS
 Location/Qualifiers
 1..3050
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
ORIGIN
 Query Match 1.9%; Score 62.4; DB 2; Length 3050;
 Best Local Similarity 51.8%; Pred. No. 3.5e-06;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757
 DB 2280 GAAGTGGCAATATCCAGCTGGAAACAAAGTTGATACAGACATTACACACCCATATGAGT 2339
 QY 2758 TCGATTTTCTTGGCATCTCATCTGTTGGTGTCTTGGTACATCTCGTCCAGGACATTACA 2817
 DB 2340 TCGATTTTTCCTCTGTAGCCATGCTGGAATACAGGTACCACTCGTCTTCACACTATC 2399
 QY 2818 CTGTTATGTATGACGATAAAGGAATGAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
 DB 2400 ATGTTTTATGGGATGATAAAGTCTTTTACTGCAGATGAAGTCTCAGCTGTCTTAACCTAC 2459
 QY 2878 TTGCTTTTCTCTCTGCTAGATGTGGAAGCCATCTCGTTGCTGTTCGGGTTCAATTATG 2937
 DB 2460 TCTGCCACACTTACGTACGCTGTACAGCATCTGTTCTTATACCTTCGACCAGCGTATTATG 2519
 QY 2938 CTCATTATCATGTGGAAGGCGAAAGAGCTT 2969
 DB 2520 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2551

Search completed: July 6, 2006, 02:44:28
 Job time : 17500 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 21:40:42 ; Search time 577 Seconds
(without alignments)
10464.589 Million cell updates/sec

Title: US-10-645-746-2

Perfect score: 3227

Sequence: 1 cagccacaaagtcatgaac.....tttaaaaaaaaaaaaaa 3227

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /EMC_Celerra_SID33/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC_Celerra_SID33/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC_Celerra_SID33/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC_Celerra_SID33/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC_Celerra_SID33/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC_Celerra_SID33/ptodata/2/ina/H COMB.seq:*
- 7: /EMC_Celerra_SID33/ptodata/2/ina/PCTUS COMB.seq:*
- 8: /EMC_Celerra_SID33/ptodata/2/ina/PP COMB.seq:*
- 9: /EMC_Celerra_SID33/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC_Celerra_SID33/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.4	1.9	734	4	US-09-297-648-2528 Sequence 2528, Ap
2	62	1.9	7218	2	US-08-232-463-14 Sequence 14, Appl
3	61.4	1.9	771	3	US-09-248-796A-6819 Sequence 6819, Ap
4	54.2	1.7	1141	3	US-09-806-708B-22 Sequence 22, Appl
5	49.6	1.5	1141	3	US-09-806-708B-22 Sequence 22, Appl
6	47	1.5	2910	3	US-09-533-029-105 Sequence 105, App
7	45.6	1.4	2328	3	US-10-043-774B-1 Sequence 1, Appli
8	45.6	1.4	3472	3	US-09-873-737A-5 Sequence 5, Appli
9	45.2	1.4	300	4	US-09-297-648-1917 Sequence 1917, Ap
10	44.2	1.4	7478	3	US-09-949-016-948 Sequence 948, App
11	44.2	1.4	7478	3	US-09-949-016-948 Sequence 948, App
12	43	1.3	832	3	US-09-621-976-2813 Sequence 2813, Ap
13	42.8	1.3	723	3	US-09-270-767-4630 Sequence 4630, Ap
14	42.8	1.3	723	3	US-09-270-767-19912 Sequence 19912, A
15	42.8	1.3	5838	3	US-08-956-1718-66 Sequence 66, Appl
16	42.8	1.3	5838	3	US-08-956-1718-66 Sequence 66, Appl
17	42.6	1.3	601	3	US-08-781-986A-66 Sequence 66, Appl
18	42.6	1.3	601	3	US-09-949-016-197722 Sequence 197722, Sequence 197723,
19	42.6	1.3	85665	3	US-09-949-016-197723 Sequence 17345, A
20	42	1.3	19124	2	US-08-487-826B-13 Sequence 13, Appl
21	41.6	1.3	1875	7	PCT-US96-05320A-997 Sequence 997, App
22	41.6	1.3	1830121	3	US-09-557-884-1 Sequence 1, Appli
23	41.6	1.3	1830121	3	US-09-643-990A-1 Sequence 1, Appli

C 24	41.6	1.3	1830121	3	US-10-158-865-1	Sequence 1, Appli
C 25	41.4	1.3	11485	3	US-09-410-464-9	Sequence 9, Appli
C 26	41.4	1.3	114426	3	US-09-949-016-15078	Sequence 15078, A
C 27	41.4	1.3	163317	4	US-09-531-120-212	Sequence 212, App
C 28	41.2	1.3	601	3	US-09-949-016-196507	Sequence 196507, A
C 29	41.2	1.3	451924	3	US-09-949-016-12896	Sequence 12896, A
C 30	41.2	1.3	451925	3	US-09-949-016-17305	Sequence 17305, A
C 31	40.8	1.3	1809	4	US-09-880-107-3376	Sequence 3376, Ap
C 32	40.6	1.3	247299	3	US-09-949-016-17590	Sequence 17590, A
C 33	40.4	1.3	2820	7	PCT-US93-11725-1	Sequence 1, Appli
C 34	40.4	1.3	47375	3	US-09-949-016-15420	Sequence 15420, A
C 35	40.4	1.3	168575	3	US-09-426-290-1	Sequence 1, Appli
C 36	39.6	1.2	1664976	3	US-08-316-421B-1	Sequence 1, Appli
C 37	39.6	1.2	1664976	3	US-09-692-570-1	Sequence 119957, Sequence 119957,
C 38	39.4	1.2	601	3	US-09-949-016-119957	Sequence 8910, Ap
C 39	39.4	1.2	1680	3	US-09-248-796A-8910	Sequence 11950, A
C 40	39.4	1.2	121427	3	US-09-949-016-11950	Sequence 13230, A
C 41	39.4	1.2	121433	3	US-09-949-016-13230	Sequence 13230, A
C 42	39.4	1.2	194889	3	US-09-949-016-15654	Sequence 15654, A
C 43	39.2	1.2	832	3	US-09-621-976-2813	Sequence 2813, Ap
C 44	39	1.2	381	3	US-09-270-767-408	Sequence 408, App
C 45	39	1.2	381	3	US-09-270-767-15690	Sequence 15690, A

ALIGNMENTS

RESULT 1

US-09-297-648-2528
Sequence 2528, Application US/09297648
Patent No. 6964868

GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.

APPLICANT: Escobedo, Jaime

APPLICANT: Innis, Michael A.

APPLICANT: Garcia, Pablo Dominquez

APPLICANT: Sudduth-Klinger, Julie

APPLICANT: Reinhard, Christoph

APPLICANT: Giese, Klaus

APPLICANT: Randazzo, Filippo

APPLICANT: Kennedy, Giulia C.

APPLICANT: Pot, David

APPLICANT: Kassan, Altaf

APPLICANT: Lamson, George

APPLICANT: Drmanac, Radoje

APPLICANT: Crkvenjakov, Radomir

APPLICANT: Dickson, Mark

APPLICANT: Drmanac, Snezana

APPLICANT: Labat, Ivan

APPLICANT: Leshkowitz, Dena

APPLICANT: Kita, David

APPLICANT: Garcia, Veronica

APPLICANT: Jones, William Lee

APPLICANT: Strache-Crain, Birgit

TITLE OF INVENTION: No. 6364868el Human Genes and Gene Expression

FILE REFERENCE: 2300-1481

CURRENT APPLICATION NUMBER: US/09/297,648

PRIOR FILING DATE: 2000-03-10

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

PRIOR FILING DATE: 1998-01-28

```

; SEQ ID NO 2528
; LENGTH: 734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(734)
; OTHER INFORMATION: n = A,T,C or G
US-09-297-648-2528

Query Match
Best Local Similarity 1.9%; Score 62.4; DB 4; Length 734;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATGTGGAACCATCATCGGACAACTGTGGATAAACTTATCGTTTGAATACAAAT 2757
Db 209 GAAGTGGCAATATCCAGCTGGGAACACAGTGTGATACAGACATTACACCCCATATGAGT 268
QY 2758 TCGATTTTCTTGGCATCTCATCATGTGTCTTGGTATCATCTCGTCCAGGACATTACA 2817
Db 269 TCGATTTTACCTCTGTAGCCATGCTGGAATACAGGTTACAGTCTGCTTACACTATC 328
QY 2818 CTGTATGATGACGATGAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 329 ATGTTTATGGGATGATAACTGCTTTACTGCAGATGAATCTCAGCTGCTAACTTACCAGC 388
QY 2878 TTGCTTTTCTCTGTAGATGTCGAAACCCATCTCGTTGCTGTTCGGTTTCATTATG 2937
Db 389 TGTGCCACACTTACGTACCTGTACACGATCTGTTTCTATACCTGCACCAGCGTATTATG 448
QY 2938 CTCATTATCATGTGAAAAAGCGAAAGAGCTT 2969
Db 449 CTCACCTGTAGCATTTAGACGCCAGATATCAT 480

```

```

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109

```

```

; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match
Best Local Similarity 1.9%; Score 62; DB 2; Length 7218;
Matches 38; Conservative 223; Mismatches 183; Indels 0; Gaps 0;

QY 2240 TCCAGGTGGAATCTATCTATCGAAATATATGTTGACTCAAGAAGAAATGCTGCCCGGTGA 2299
Db 1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378
QY 2300 CGGTCCAGTGGCTCATGCGAGGGAAGACACATATTTTGGGAAGCAAGTTGCTGAAATT 2359
Db 1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318
QY 2360 GCTCAGAGAATTCGAGAGAAAACAACGACAATCGAGCACCGCATATTTGATCTTATCG 2419
Db 1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258
QY 2420 AGACGGATTAGCGATTGGAGATGCTAGCTGTAGTCAATGATGAGCTTCGATCTTTAA 2479
Db 1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198
QY 2480 AAGCGAAGTAAACAATTCATGTCGAGCGGATGAGAGATCCAGAGCCGAAGTACAC 2539
Db 1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138
QY 2540 GTTCATTGTGATTCAGAAAAGACACAATACACATGCTTCGAGAAGATTCGAGAAGATGAA 2599
Db 1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078
QY 2600 GCCAGTGTCAATAAAGATCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2659
Db 1077 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1018
QY 2660 ACAATGGGAGGAGGATATGAAAGA 2683
Db 1017 CGTATGGCAACGAGGAAAATA 994

```

```

RESULT 3
US-09-248-796A-6819
; Sequence 6819, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6819
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6819

```

```

Query Match
Best Local Similarity 1.9%; Score 61.4; DB 3; Length 771;
Matches 146; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```


Db 2066 ATGACCTTTTATGTCGAGCCAGGCTGTGAGAGTGGTAGTGTCTCCACACATTACA 2125
Qy 2818 CTGTTATGATGACGATTAAGAAATGAGCCAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 2126 ATGTCATCTATGACAAACAGCGGCTGAAGCCAGACCATACAGCGCTTGACCTACAGC 2185
Qy 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTGTCCTGCTGCTGCTTCAATATG 2937
Db 2186 TGTGCCACATCTATTACAACTGGCCAGGTGTCATTCGTGTTCTCTGCTCTCTTGGCCAGTACG 2245
Qy 2938 CTCA 2941
Db 2246 CCCA 2249

RESULT 8
US-09-873-737A-5
; Sequence 5, Application US/09873737A
; Patent No. 6723534
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED p1v1 FAMILY GENES AND GENE
; TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6723534 180-104/2
; CURRENT APPLICATION NUMBER: US/09/873,737A
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110,901
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (167)..(2749)
; NAME/KEY: misc feature
; LOCATION: (392)
; OTHER INFORMATION: n-t or a, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (1073)
; OTHER INFORMATION: n-c or a, Xaa=Leu or Ile
; NAME/KEY: misc feature
; LOCATION: (2365)
; OTHER INFORMATION: n-c or a, Xaa=Leu or Ile
US-09-873-737A-5

Query Match 1.4%; Score 45.6; DB 3; Length 3472;
Best Local Similarity 49.2%; Pred. No. 0.3;
Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 2698 GAATTGTGAACCATCATCCGGAACAACTGTGGATAAACTTATGTTTCGAAATACAAAT 2757
Db 2430 GACTTCAGAATCCACTTCTCTGGAACAGTATTGATGTAGAGTTTACCAGACCAGAAATGGT 2489
Qy 2758 TCGATTTTCTTGGCATCTCATCATGCTGCTTGGTATCATCTCGTCCAGGACATTACA 2817
Db 2490 ATGACCTTTTATCGTGAGCCAGGCTGTGAGAAGTGGTAGTGTCTTCCACACATTACA 2549
Qy 2818 CTGTTATGATGACGATTAAGAAATGAGCCAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 2550 ATGTCATCTATGACAAACAGCGGCTGAAGCCAGACCATACAGCGCTTGACCTACAGC 2609
Qy 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTGTCCTGCTGCTGCTTCAATATG 2937
Db 2610 TGTGCCACATCTATTACAACTGGCCAGGTGTCATTCGTGTTCTCTGCTCTCTTGGCCAGTACG 2669
Qy 2938 CTCA 2941

Db 2670 CCCA 2673

RESULT 9
US-09-297-648-1917
; Sequence 1917, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassar, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; TITLE OF INVENTION: Products II
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1917
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-1917

Query Match 1.4%; Score 45.2; DB 4; Length 300;
Best Local Similarity 55.8%; Pred. No. 0.13;
Matches 86; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 2698 GAATTGTGAACCATCATCCGGAACAACTGTGGATAAACTTATGTTTCGAAATACAAAT 2757
Db 147 GAATGGCAATATCCAGCTGGAAACACAGTTGTATACAGACATTACACACCCCATATGAT 206
Qy 2758 TCGATTTTCTTGGCATCTCATCATGCTGCTTGGTATCATCTCGTCCAGGACATTACA 2817
Db 207 TCGATTTTACCTTCTGTAGCCATGCTGGAAATACAGGGTACAGGCTCTCTTCACTATC 266
Qy 2818 CTGTTATGATGACGATTAAGAAATGAGCCAGATGAAGTCTATAAAATGACCTACGGAC 2851
Db 267 ATGTTTATGGGATGATAACTGCTTTACTGCACA 300

RESULT 10
US-09-949-016-948
; Sequence 948, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-948

Query Match 1.4%; Score 44.2; DB 3; Length 7478;
Best Local Similarity 49.0%; Pred. No. 0.95;
Matches 118; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-948

RESULT 11
US-09-949-016-4453
; Sequence 4453, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4453
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4453

US-09-949-016-4453
Query Match 1.4%; Score 44.2; DB 3; Length 7478;
Best Local Similarity 49.0%; Pred. No. 0.95;
Matches 118; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4453

RESULT 12
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.prm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 1.3%; Score 43; DB 3; Length 832;
Best Local Similarity 12.5%; Pred. No. 0.73;
Matches 46; Conservative 167; Mismatches 152; Indels 2; Gaps 1;
GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4453
; LENGTH: 7478
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4453

Qy	622	CTTTGTGACGACGAGATTAAAGTACAATTGCGGAAAAATTTTGTGTACGATAATAATT	681
Db	311	WVCWARYRSTGTGTRASMMWRWYTTMMKKWKYAWADAAWRWWAWWWAARRACAAAATA	370
Qy	682	CAATTCT	688
Db	371	TAATTAT	377

```

RESULT 13
US-09-270-767-4630/c
; Sequence 4630, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4630
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-4630

```

```

RESULT 14
US-09-270-767-19912/c
; Sequence 19912, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLS OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19912
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-19912

Query Match 1.3% Score 42.8; DB 3; Length 723;
Best Local Similarity 48.0%; Pred. No. 0.78;

```

Matches 122; Conservative	0;	Mismatches 132; Indels	0;	Gaps	0;
Qy	1119	AAATCGGACAGAGGAATAATCAAAAAAGTAGCATACATCTGTTCAAAAATCTATGAG	1178		
Db	722	AAAATGGAAAAATGCGATGAAAGCAAAAAGCTTAAGAAAACTTGAACAGGCTTTAGCACTTTG	663		
Qy	1179	GAAACAAAAAGTTTCATTGAGTTTCCCGACCTACCACCTAGTCAAAGTTAAAAAGTGGAGCA	1238		
Db	662	GCAACCGAAGAAACGATCTCAATTAATAAAAAAATGAGATTCCCAATTAAAGAAAAA	603		
Qy	1239	AAAGTAATACGCTGTACCAATGGAAACATCTTGAAGTTTCATGAGAAGCCACAAAGATACAAG	1298		
Db	602	AAACAAATATAAGAAATGGAATAAATGAATAAACAATTTATAAGAAATCATACAAT	543		
Qy	1299	AATCGAATTGATCTCGTGATGCAAGACAAAGTTTCTAAAGCGAGCTACACGAAAACTCAC	1358		
Db	542	AGGCCAATACCAACGTTTTTAGCAATCTAAGTCTAAGAAGAAAACTAATCGAAATTTTCTT	483		
Qy	1359	GACTACAAAGAAA	1372		
Db	482	TACAACAAAATAA	469		

RESULT 15
 US-08-956-171E-66/c
 ; Sequence 66, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; Gil H. Choi
 ; Patrick S. Dillon
 ; Craig A. Rosen
 ; Steven C. Barash
 ; Michael R. Fannon
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/956,171E
 ; FILING DATE: 20-Oct-1997
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/009,861
 ; FILING DATE: January 5, 1996
 ; APPLICATION NUMBER: 08/781,986
 ; FILING DATE: January 3, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mark J. Hyman
 ; REGISTRATION NUMBER: 46,789
 ; REFERENCE/DOCKET NUMBER: PB248P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (240) 314-1224
 ; TELEFAX: (301) 309-8439
 ; INFORMATION FOR SEQ ID NO: 66:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5838 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 66:
 US-08-956-171E-66

```
Query Match      1.3%; Score 42.8; DB 3; Length 5838;
Best Local Similarity 50.0%; Pred. No. 1.9; Mismatches 0; Gaps 0;
Matches 107; Conservative 0; Indels 107; Indels 0; Gaps 0;

Qy 1361 CTACAAAGAAAAATACCCCTAAAAATGCTGAAAGAAATTTGGATTTCTCTTCTGAAAGAGCTAAA 1420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 CCACATCCAGCTAACAAATATGTTGGCCCAACAATTTGAAACCAATTTTTTCATTACTATC 545

Qy 1421 TTTTGTGAAAGATTTGGATTTATGCTCCAAACTTCAGATGATCGAATGTCCAGGAAAGGT 1480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 TCCCTAAAAATCAATATCTATTTATTTTATCCTTATTATGATATTATTTAATGTATAAA 485

Qy 1481 TTTGAAAGAGCCCAATCCTTGTGAATAGTGTAAATGAACAAATTTAAAATGACACCAGTGAT 1540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 GTTGAANTGCACAAACCGTTCACCTAGGAGGTATATAAGATGAGCAGAAATACCAAGTGCA 425

Qy 1541 TCGTGGATTTCAAGAAAAACAATTTGAATGTGGTT 1574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 TCGTATTTTGAACAAAACATTTAGATGCGATT 391
```

Search completed: July 5, 2006, 22:10:06
Job time : 580 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 21:15:33 ; Search time 1721 Seconds
(without alignments)
13073.472 Million cell updates/sec

Title: US-10-645-746-2
Perfect score: 3227
Sequence: 1 cagccacaagtgtgaac.....tttaaaaaaaaaaaaaa 3227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3227	100.0	3227	4	AAS03283 C. elegans
2	2497.4	77.4	3709	4	AAS03282 C. elegans
3	63.8	2.0	2827	9	ADb17459 Soybean p
4	63.8	2.0	2827	14	Aec75683 Soybean Z
5	62.4	1.9	490	9	ACH23411 Human adu
6	62.4	1.9	734	2	AAZ15059 Human gen
7	62.4	1.9	2568	12	ADJ94706 Human euk
8	62.4	1.9	2914	4	AAS33425 Human pol
9	62.4	1.9	2914	8	ABX13625 Human cyt
10	62.4	1.9	3050	4	AAH16058 Human cdn
11	62.4	1.9	3050	12	ADH77114 Human PAZ
12	62.4	1.9	3325	8	AAD55860 Human nuc
13	62.4	1.9	3736	8	ABX34752 Human mdd
14	61.8	1.9	2679	13	ADQ80688 Arabidops
15	60	1.9	1501	9	ADb17461 Soybean p
16	60	1.9	1501	14	AEC75685 Soybean Z
17	57	1.8	441	9	ACH28746 Human adu
18	56.4	1.7	2772	12	ADJ94705 Human euk

19	54.8	1.7	2951	4	AAH24562	Aah24562 Translati
20	54	1.7	2000	11	ACL35363	ACL35363 Rice stre
21	53	1.6	3520	13	ADQ80702	Adq80702 Arabidops
22	52.8	1.6	3096	9	ADb17463	Adb17463 Corn post
23	52.8	1.6	3096	14	AEC75687	Aec75687 Corn Argo
24	51.4	1.6	1565	9	ADb17505	Adb17505 Wheat pos
25	51.4	1.6	1565	14	AEC75729	Aec75729 Wheat Arg
26	51	1.6	1158	3	AAC50998	Aac50998 Arabidops
27	51	1.6	3120	13	ADQ80694	Adq80694 Arabidops
28	51	1.6	3147	12	ADi43622	Adi43622 plant tra
29	51	1.6	3195	3	AAC50992	Aac50992 Arabidops
30	51	1.6	3326	3	AAC47079	Aac47079 Arabidops
31	50.2	1.6	1238	9	ADb17487	Adb17487 Rice post
32	50.2	1.6	1238	14	AEC75711	Aec75711 Rice Argo
33	49.8	1.5	2960	13	ADQ80704	Adq80704 Arabidops
34	49.8	1.5	2973	12	ADQ80720	Adq80720 Transcrip
35	49.8	1.5	3008	12	ADQ63642	Ado63642 Transcrip
36	49.6	1.5	2622	13	ADQ80690	Adq80690 Arabidops
37	48.8	1.5	2775	13	ADQ80686	Adq80686 Arabidops
38	48.8	1.5	2793	6	ABZ12589	Abz12589 Arabidops
39	48.8	1.5	2793	10	ADb31434	Adb31434 Plant vie
40	48.8	1.5	2793	12	ADi44098	Adi44098 Plant tra
41	48.6	1.5	2000	8	ADA71938	Ada71938 Rice gene
42	47.8	1.5	2960	13	ADQ80696	Adq80696 Arabidops
43	47.8	1.5	2967	3	AAC50603	Aac50603 Arabidops
44	47.8	1.5	2967	6	ABZ12627	Abz12627 Arabidops
45	47.8	1.5	3071	14	ADY65087	Ady65087 S. manson

ALIGNMENTS

RESULT 1

AAS03283

ID AAS03283 standard; cDNA; 3227 BP.

XX AAS03283;

AC AAS03283;

DT 07-SEP-2001 (first entry)

XX C. elegans cDNA encoding RNA interference pathway protein RDE-1.

DE RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;

XX double-stranded RNA-dependent gene silencing; ss.

XX Caenorhabditis elegans.

OS Location/Qualifiers

FT 5'UTR

FT 1..20

FT /*tag= a

FT CDS

FT 21..3083

FT /*tag= b

FT /product= "RDE-1"

FT XX

PN WO200129058-A1.

XX 26-APR-2001.

XX 13-OCT-2000; 2000WO-US028470.

XX 15-OCT-1999; 99US-0159776P.

XX 30-MAR-2000; 2000US-0193218P.

XX (UYMA-) UNIV MASSACHUSETTS.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX Mello CC, Fire A, Tabara H, Grishok A;

XX WPI; 2001-316239/33.

XX P-PSDB; AAU01856.

XX Novel RNA interference pathway genes and their protein products involved

XX in mediation of genetic interference, useful for modulating and studying

regulation of RNA interference pathway.

Claim 1: Fig 6: 76pp: English.

The sequence encodes the RNA interference (RNAi) pathway protein RDE-1. RDE-1 and RDE-4 are involved in the pathway mediating double-stranded RNA-dependent gene silencing (genetic interference) RDE-1 and RDE-4 protein is useful for preparing an RNAi agent, by incubating a dsRNA in the presence of the proteins. The prepared RNAi agents can be used as sequence-specific interfering agents for targeted genetic interferences. The nucleic acids are useful for studying the regulation of RNAi pathway and to generate knockout strains of animals such as *C.elegans*. RDE-1 and RDE-4 genes and their products are useful for modulating RNAi pathway activity. The polypeptides are useful for generating and testing antibodies specific for the polypeptides which are useful for studying the RNAi pathway in *C.elegans* and other organisms. RNAi pathway genes are useful for mediating specific processes, e.g. a gene that mediates dsRNA uptake by the cells is useful for transporting other RNAs into cells or for facilitating entry of agents such as drugs into cells. RNAi pathway mutant strains (rde-1, rde-4 mutants) are useful in genetic screens to identify additional RNAi pathway components

Sequence 3227 BP; 1145 A; 575 C; 666 G; 841 T; 0 U; 0 Other;

Query Match 100.0%: Score 3227: DB 4: Length 3227;

Best Local Similarity 100.0%; Pred. NO. 0;
Matches 3227; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 CAGCCACAAAGTGATGAAACATGTCTCGAATTTCCCGAATTTGGAAAAGGATTTTATC 60

1 CAGCCACAAAGTGATGAAACATGTCTCCCGAATTTCCCGAATTGGAAAAAGGATTTTATC 60

61 GTCATTCTCTCGATCCGGAGATGAAATGGCTTGGAGGCCCACTGGTAAATGCGACGGCA 120

61 GTCATTCTCTCGATCCGGAGATGAAATGGCTTGCAGGCCCACTGGTAAATGCGACGGCA 120

121 AATTCTATGAGAGAAAGTACTTCTTTTGGTAAATTGGTTCAAGTTCTCCAGCAAAATTT 180

121 AATTCTATGAGAGAAAGTACTTCTTTTGGTAAATTGGTTCAAGTCTCCAGCAAAATT 180

181 ACGATCGGGAATACTACGAGTGAAGTGAAATGACAAAGGAAGTATTGAATAGAAAC 240

181 ACGATCGGGAATACTACGAGTATGAAGTGAAAATGACAAAGGAAGTATTGAATAGAAAAC 240

241 CAGGAAACCTTTCCCAAAAGACAGAAATTCCAATTCCTCGATCGTGCAAACTCTTCT 300

241 CAGGAAACCTTTCCCAAAGACAGAAATTCCAATTCCTCGTGCAGAACTCTCT 300

301 GGCAACATCTTCGGCATGAGAAGCAGACAGATTTTATTCCTGAAGACTATGTTTGTG 360

301 GGCAACATCTTCGGCATGAGAAGCAGACAGATTTATTCTCGAAGACTATGTTTGTG 360

361 ATGAAAAGGACACTGTTTATAGTGTTTGTCTGACTGAACACTGTCAATCAAAAATGCTGG 420

361 ATGAAAGGACACTGTTTATAGTGTTCGACTGAACACTGTCAATCAAAATGCTGG 420

421 TTTCCGAGAAAGTAGTAAATAAAGGATTCGGAGAAAAGATGAAAAGGATTCGGAGAAA 480

421 TTTCCGAGAAAGTAGTAAAAAAGGATTCGGAGAAAAAAGATGAAAAGGATTTCGGAGAAA 480

481 AAATCTTATACACAATGATACTTACCTATCGTAAAAATTTCACTGAACCTTAGTCGAG 540

481 AAATCTTATACACAAATGATACCTTACCTATCGTAAAAAATTTACCTGAACCTTTAGTCGAG 540

541 AAAATCCGAAAGACGAGCGAATCCGAGTTACAAATTCCTGAAGAAATGTTATGA 600

541 AAAATCCGGAAAGAGACGAAGCGAAATCGGAGTTACAAATTCCTGGAAGAAATGTTATGA 600

601 CCCAGAAAGTTCGCTACGGCCCTTTTGTGAACGAGGAGATTAAAGTACAAATTCGCGAAA 660

601 CCCAGAAAGTTCGCTACGCGCCTTTTGTGTAACGAGGAGATTAAAGTACAAATTCGCGAAA 660

661 ATTTGTGTACGATAATAATTCAATTCTGCGAGTTCCTGAATCGTTTCACGATCCAAACA 720

Db	1741	TGTACGACGCGACGAAATAATGAATATGCTTCTCAAAATAATGTTACACTAAATATACCGGAA	1800
Qy	1801	TCGGTAGATTTGAAATAGCCGCAACAGAAAGCGAAAGAAATATGTTTGAACGCTCTTCCCGATA	1860
Db	1801	TCGGTAGATTTGAAATAGCCGCAACAGAAAGCGAAGAAATATGTTTGAACGCTCTTCCCGATA	1860
Qy	1861	AAGAACAAAAGTCTTAATGTTTCAATTATCTTCCAAACGACAACTGAATGCTTTACGGTT	1920
Db	1861	AAGAACAAAAGTCTTAATGTTTCAATTATCTTCCAAACGACAACTGAATGCTTTACGGTT	1920
Qy	1921	TTGTGAAACATTTATTCGCATCACACCATCGGTGTAGCTTAATCAGCATATTACTTCTGAAA	1980
Db	1921	TTGTGAAACATTTATTCGCATCACACCATCGGTGTAGCTTAATCAGCATATTACTTCTGAAA	1980
Qy	1981	CAGTCACAAAAGCTTTGGCATCACTTAAGGCCAGGAAAGGATCAAAACGAAATTTCTATC	2040
Db	1981	CAGTCACAAAAGCTTTGGCATCACTTAAGGCCAGGAAAGGATCAAAACGAAATTTCTATC	2040
Qy	2041	AAATTCGATTTGAAATCAACCGGAAATTTAGGAGGTAATTAACGAGGCTTGACTGTGCAG	2100
Db	2041	AAATTCGATTTGAAATCAACCGGAAATTTAGGAGGTAATTAACGAGGCTTGACTGTGCAG	2100
Qy	2101	AAATTCGAGAAATATCACCGAAGAAAGAAACGCGGAAACAAATGCCATTAACTATGT	2160
Db	2101	AAATTCGAGAAATATCACCGAAGAAAGAAACGCGGAAACAAATGCCATTAACTATGT	2160
Qy	2161	ATGTTGGAATTTGATGTAAGTAACTCATCCAACTCTCTACAGTGGAAATGATTTATCTATACGG	2220
Db	2161	ATGTTGGAATTTGATGTAAGTAACTCATCCAACTCTCTACAGTGGAAATGATTTATCTATACGG	2220
Qy	2221	CTGTAGTAGCGAGTATCAATCCAGGTGGAACTATCTATCGAAATPATGATTTGTGACTCAAG	2280
Db	2221	CTGTAGTAGCGAGTATCAATCCAGGTGGAACTATCTATCGAAATPATGATTTGTGACTCAAG	2280
Qy	2281	AAGAAATGTCGTCGCGGTGAGCGGTGCAGTGGCTCATGACGCGGAAAGAAACAGATATTTTGG	2340
Db	2281	AAGAAATGTCGTCGCGGTGAGCGGTGCAGTGGCTCATGACGCGGAAAGAAACAGATATTTTGG	2340
Qy	2341	AAGCAAAGTTCGTCGAAATTTGCTCAGAGAAATTCGACAGAAACAAACGACAAATCGAGCACCAG	2400
Db	2341	AAGCAAAGTTCGTCGAAATTTGCTCAGAGAAATTCGACAGAAACAAACGACAAATCGAGCACCAG	2400
Qy	2401	CGCATATTTGTAGTCTATTCGAGACCGAGTTAGCGAATTCGAGAGTGCTACGTTTAGTCAATG	2460
Db	2401	CGCATATTTGTAGTCTATTCGAGACCGAGTTAGCGAATTCGAGAGTGCTACGTTTAGTCAATG	2460
Qy	2461	ATGAGCTTCGATCTTTAAAAGCCAGTAAAACAAATTCATGTCGGAAACGGGATGGAGAAG	2520
Db	2461	ATGAGCTTCGATCTTTAAAAGCCAGTAAAACAAATTCATGTCGGAAACGGGATGGAGAAG	2520
Qy	2521	ATCCAGAGCCGAGTACACGTTTCATTCGTGATTCAGAAAGACACAAATACACAGATTCGCTTC	2580
Db	2521	ATCCAGAGCCGAGTACACGTTTCATTCGTGATTCAGAAAGACACAAATACACAGATTCGCTTC	2580
Qy	2581	GAAGAAATGGAAAAAGATAAGCCAGTGGTCAATTAAGATCTTTACTCTGCTGAAAACAGATG	2640
Db	2581	GAAGAAATGGAAAAAGATAAGCCAGTGGTCAATTAAGATCTTTACTCTGCTGAAAACAGATG	2640
Qy	2641	TCGCTGTGCTGCTGTTTAAACAAATGGGAGGAGGATATGAAAGAAAGCAAAAGAACTGGAA	2700
Db	2641	TCGCTGTGCTGCTGTTTAAACAAATGGGAGGAGGATATGAAAGAAAGCAAAAGAACTGGAA	2700
Qy	2701	TTGTGAAACCCATCATCCGGAAACACTGTGGATAAATTCGTTTCGAAATACAAATTCG	2760
Db	2701	TTGTGAAACCCATCATCCGGAAACACTGTGGATAAATTCGTTTCGAAATACAAATTCG	2760
Qy	2761	ATTTTTTTCTTGGCATCTCATCATGCTGCTCTTGGTACATCTCGTCCAGGACATTAACACTG	2820
Db	2761	ATTTTTTTCTTGGCATCTCATCATGCTGCTCTTGGTACATCTCGTCCAGGACATTAACACTG	2820
Qy	2821	TTATGTATGACGATAAAGGAATGAGCCAGCAAGATGAAGTCTATAAATGACCTACGACTTG	2880
Db	2821	TTATGTATGACGATAAAGGAATGAGCCAGCAAGTGAAGTCTATAAATGACCTACGACTTG	2880

Qy	2881	CTTTTCTCTCTGCTAGATGTCGAAAA	CCCAATCTCTGTTGCCGTGTTCCGGTTCAATTAGCTC	2940
Db	2881	CTTTTCTCTCTGCTAGATGTCGAAAA	CCCAATCTCTGTTGCCGTGTTCCGGTTCAATTAGCTC	2940
Qy	2941	ATTTTATCATGTGAAAAAGCGAAAGAG	CTTTTATCGAACTTACAAGGAACATTACATCGGTG	3000
Db	2941	ATTTTATCATGTGAAAAAGCGAAAGAG	CTTTTATCGAACTTACAAGGAACATTACATCGGTG	3000
Qy	3001	ACTATGCACAGCCAGCACTCGACAGAAATG	GAACATTTTCTCCAAACTTAAAGTGAAGT	3060
Db	3001	ACTATGCACAGCCAGCACTCGACAGAAATG	GAACATTTTCTCCAAACTTAAAGTGAAGT	3060
Qy	3061	ACCTCGGAATGCTGTTGCGCATAA	CATTTTGC AAAAGTGTCCCGCTTCAATCAAAATTTT	3120
Db	3061	ACCTCGGAATGCTGTTGCGCATAA	CATTTTGC AAAAGTGTCCCGCTTCAATCAAAATTTT	3120
Qy	3121	TCAATTCGTAGATATGTACTTAC	TTTTTTTTTAAAGCCCGTTTCAAAAATTCATTCCATG	3180
Db	3121	TCAATTCGTAGATATGTACTTAC	TTTTTTTTTAAAGCCCGTTTCAAAAATTCATTCCATG	3180
Qy	3181	ACTAACGTTTTTCATAAAATTA	CTTGAAATTTAAAAAATAAAAAAAAAA	3227
Db	3181	ACTAACGTTTTTCATAAAATTA	CTTGAAATTTAAAAAATAAAAAAAAAA	3227

РЕСПУБЛИКА

RESULT 2
AAG03303

AAS03282
ID AAS03282 standard; DNA; 3709 BP.

XX

AC AAS03282;

1

DT 07-SEP-2001 (first entry)

DE C. elegans genomic DNA encoding RNA interference pathway protein RDE-1.

XX

KW RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;

KW double-stranded RNA-dependent

XX
XXXXXXXXXXXX

OS *Caenorhabditis elegans*.XX
XXXXXXXXXXXXX

Key	Location/Qualifiers
101	101
102	102
103	103
104	104
105	105
106	106
107	107
108	108
109	109
110	110
111	111
112	112
113	113
114	114
115	115
116	116
117	117
118	118
119	119
120	120
121	121
122	122
123	123
124	124
125	125
126	126
127	127
128	128
129	129
130	130
131	131
132	132
133	133
134	134
135	135
136	136
137	137
138	138
139	139
140	140
141	141
142	142
143	143
144	144
145	145
146	146
147	147
148	148
149	149
150	150
151	151
152	152
153	153
154	154
155	155
156	156
157	157
158	158
159	159
160	160
161	161
162	162
163	163
164	164
165	165
166	166
167	167
168	168
169	169
170	170
171	171
172	172
173	173
174	174
175	175
176	176
177	177
178	178
179	179
180	180
181	181
182	182
183	183
184	184
185	185
186	186
187	187
188	188
189	189
190	190
191	191
192	192
193	193
194	194
195	195
196	196
197	197
198	198
199	199
200	200
201	201
202	202
203	203
204	204
205	205
206	206
207	207
208	208
209	209
210	210
211	211
212	212
213	213
214	214
215	215
216	216
217	217
218	218
219	219
220	220
221	221
222	222
223	223
224	224
225	225
226	226
227	227
228	228
229	229
230	230
231	231
232	232
233	233
234	234
235	235
236	236
237	237
238	238
239	239
240	240
241	241
242	242
243	243
244	244
245	245
246	246
247	247
248	248
249	249
250	250
251	251
252	252
253	253
254	254
255	255
256	256
257	257
258	258
259	259
260	260
261	261
262	262
2	

21	3582
----	------

[illegible]

```

21  IF
    /product = "PG-1"
    /tag = a

```

```
FT /product="KDS-I"
FT /note="This section
```

Y F I F I

XX
PN
W03001300ER-21

FN
yy
W0200129038-AT.

[illegible]

FD 26-APR-2001.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PF 13-OCT-2000; 2000WO-US028470.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PR 15-OCT-1999; 99US-0159776P.

PR 30-MAR-2000; 2000US-0193218P.

XX

PA (UYMA-) UNIV MASSACHUSETTS.

PA (CARN-) CARNEGIE INST WASHINGTON.

XX

CC and RDE-4 protein is useful for preparing an RNAi agent, by incubating a
 CC dsRNA in the presence of the proteins. The prepared RNAi agents can be
 CC used as sequence-specific interfering agents for targeted genetic
 CC interferences. The nucleic acids are useful for studying the regulation
 CC of RNAi pathway and to generate knockout strains of animals such as
 CC C.elegans. RDE-1 and RDE-4 genes and their products are useful for
 CC modulating RNAi pathway activity. The polypeptides are useful for
 CC generating and testing antibodies specific for the polypeptides which are
 CC useful for studying the RNAi pathway in C.elegans and other organisms.
 CC RNAi pathway genes are useful for mediating specific processes, e.g. a
 CC gene that mediates dsRNA uptake by the cells is useful for transporting
 CC other RNAs into cells or for facilitating entry of agents such as drugs
 CC into cells. RNAi pathway mutant strains (rde-1, rde-4 mutants) are useful
 CC in genetic screens to identify additional RNAi pathway components
 XX
 SQ Sequence 3709 BP; 1308 A; 621 C; 741 G; 1039 T; 0 U; 0 Other;

Query Match 77.4%; Score 2497.4; DB 4; Length 3709;
 Best Local Similarity 85.8%; Pred. No. 0;
 Matches 3188; Conservative 0; Mismatches 16; Indels 511; Gaps 17;
 QY 1 CAGCCACAAAGTGATGAACATGCTCTCGAATTTTCCGAAATGGAAAAAGGATTTATC 60
 DB 1 CAGCCACAAAGTGATGAACATGCTCTCGAATTTTCCGAAATGGAAAAAGGATTTATC 60
 QY 61 GTCATTTCTCTCGATCC----- 76
 DB 61 GTCATTTCTCTCGATCCGGTATGATCAATTTATGACAGCTATTAAGATATATAAGTTTGATA 120
 QY 77 -----GGAGATGAATGGCTTGGAGGCCCACTGGTAAATCGACGGCAAAATTC 125
 DB 121 TTAATATTATAGGATGAATGGCTTGGAGGCCCACTGGTAAATCGACGGCAAAATTC 180
 QY 126 TATGAGAAAGATGATCTTTTGGTAAATGGTTCAGATTTCTCCAGCAAAATTTACGAT 185
 DB 181 TATGAGAAAGATGATCTTTTGGTAAATGGTTCAGATTTCTCCAGCAAAATTTACGAT 240
 QY 186 CGGGAATACACGATGATCAAGTGAAATGACAAAGGAAGTATTGAATAGAAACACGGA 245
 DB 241 CGGGAATACACGATGATCAAGTGAAATGACAAAGGAAGTATTGAATAGAAACACGGA 300
 QY 246 AAACCTTTCCCAAAAAGACAGA----- 268
 DB 301 AAACCTTTCCCAAAAAGAGAAATTCCAATGTAAGTCTTGTAATTTAGTCAAAACTAAT 360
 QY 269 ----AATTCGAATTCGCGATCGTGAACATCTTTGGCAACATCTTCGGCATGAGAGAA 325
 DB 361 TTTATTTTTCAGTCCCGATCGTCAAAACTCTTTGGCAACATCTTCGGCATGAGAGAA 420
 QY 326 GCAGACAGATTTTATTTCTCGAGACTATGTTTTTGTGAAAGGACACCTGTTTATAGTCT 385
 DB 421 GCAGACAGATTTTATTTCTCGAGACTATGTTTTTGTGAAAGGACACCTGTTTATAGTCT 480
 QY 386 TTGTGACTGAACACATGTCACATCAAAAATGCTGTTTCGGAGAAAGTAGTAAAAAGGA 445
 DB 481 TTGTGACTGAACACATGTCACATCAAAAATGCTGTTTCGGAGAAAGTAGTAAAAAGGA 540
 QY 446 TTCGGAGAAAAAGATGAAGAGATTTGGAGAAAAAATCTTTATACAAATGATACCTTAC 505
 DB 541 TTCGGAGAAAAAGATGAAGAGATTTGGAGAAAAAATCTTTATACAAATGATACCTTAC 600
 QY 506 CTATCGTAAATAATTTTCACTGAACCTTTAGTCGAGAAATCCGGAAGAGACGAGAGAGC 565
 DB 601 CTATCGTAAATAATTTTCACTGAACCTTTAGTCGAGAAATCCGGAAGAGAGAGAGAGC 660
 QY 566 GAATCGGAGTTACAAATTTCTCGA----- 588
 DB 661 GAATCGGAGTTACAAATTTCTCGAAGGTTTATGAABAAACACGCAATTAACAAACATTAG 720
 QY 589 -----AGATGTTATGACCCAGAAAGTTTCGCTACGGCCCTTTTGTGAACGAGAGATTAA 643
 DB 721 CTTTTCAGATGTTATGACCCAGAAAGTTTCGCTACGGCCCTTTTGTGAACGAGAGATTAA 780

QY 644 ACT-----ACA 649
 DB 781 AGTGTGAGTTGCAATAATAATAATCACTCACTCACTTATATATATTTAAGACA 840
 QY 650 ATTCCGCAAAATTTTGTGTACGATAATAATTTCAATTTCTGCGAGTTCTGGAATCGTTTCA 709
 DB 841 ATTCCGCAAAATTTTGTGTACGATAATAATTTCAATTTCTGCGAGTTCTGGAATCGTTTCA 900
 QY 710 CGATCCAAACAGATTGCAACATCATTAGAACTAGCAGCAAGCAAGTCAAGCATGTTTGG 769
 DB 901 CGATCCAAACAGATTGCAACATCATTAGAACTAGCAGCAAGCAAGTCAAGCATGTTTGG 960
 QY 770 AATTTCATTTGGAATCAAAAGAAATTTGTCGATGGTGAACCTGTGCTCAATTTTGGCAA- 825
 DB 961 AATTTCATTTGGAATCAAAAGAAATTTGTCGATGGTGAACCTGTGCTCAATTTTGGCAAGTAA 1020
 QY 826 -----TTGTCGATAAACTATTTC 842
 DB 1021 GTTTGAGAAACCTGCGATAAAAAAATCATGTGATTTTGTGTAAGTTTGTGATAAACTATTTC 1080
 QY 843 TACAAATGCAACCGAAATGCTCTCTGGAATTTATCTTCTTAATTTGTCGACCCCGAGTCG 902
 DB 1081 TACAATGCAACCGAAATGCTCTCTGGAATTTATCTTCTTAATTTGTCGACCCCGAGTCG 1140
 QY 903 TGTAAACGATGATGTACGAAAGATCTTAAACAAAAAATCTGATGCGCGGAAAAAATGACAAATC 962
 DB 1141 TGTAAACGATGATGTACGAAAGATCTTAAACAAAAAATCTGATGCGCGGAAAAAATGACAAATC 1200
 QY 963 AGCAAGCGCGCGGCCCAAGAAATTCGACAAATTTTGGAAAAATTTGAAGCTGAAATCGGCA 1022
 DB 1201 AGCAAGCGCGCGGCCCAAGAAATTCGACAAATTTTGGAAAAATTTGAAGCTGAAATCGGCA 1260
 QY 1023 GAAATTTGGGATAACGAA----- 1040
 DB 1261 GAAATTTGGGATAACAAATTTTGTAGTTTAAATTTATCAAACAATTAATATACAAATTTGATT 1320
 QY 1041 ---ATGTGAGATTTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGGAAAACTCT 1097
 DB 1321 TTCAGGTGAGATTTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGGAAAACTCT 1380
 QY 1098 CTTGTTTATAAGTCACTGGTAAATCGGACAGAGAGAAATGCAAAAAAGTACGATACT 1157
 DB 1381 CTTGTTTATAAGTCACTGGTAAATCGGACAGAGAGAAATGCAAAAAAGTACGATACT 1440
 QY 1158 ACATTTGTTCAAAATCTATGAGGAAACCAAAAGATTCATTGAGTTTCCCGACCTTACCACCTA 1217
 DB 1441 ACATTTGTTCAAAATCTATGAGGAAACCAAAAGATTCATTGAGTTTCCCGACCTTACCACCTA 1500
 QY 1218 GTCAAAGTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATGGAAACATCTTGAAGTTCAAT 1277
 DB 1501 GTCAAAGTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATGGAAACATCTTGAAGTTCAAT 1560
 QY 1278 GAGAGCCCAAAAGATACAGAAATCGAATTTGATCTGGTATGTCAGACAAAGTTTCTTAAG 1337
 DB 1561 GAGAGCCCAAAAGATACAGAAATCGAATTTGATCTGGTATGTCAGACAAAGTTTCTTAAG 1619
 QY 1338 CGAGCTACACGAAACCTTCCGACCTACAAAGAAATACCCCTAAATGCTGAAAGAAATTTG 1397
 DB 1620 CGAGCTACACGAAACCTTCCGACCTACAAAGAAATACCCCTAAATGCTGAAATTTGAA 1678
 QY 1398 GATTTCTCTTCTGAAAGAGCTAAATTTTGTGAAAGATTTGGATTTATGCTCCAAACTTCAG 1457
 DB 1679 GATTTCTCTTCTGAAAGAGCTAAATTTTGTGAAAGATTTGGATTTATGCTCCAAACTTCAG 1738
 QY 1458 ATGATCGAATGTCCAGGAAAGGTTTTTGAAGCCAAATGCTTGTGAAATAGTAAATGAA 1517
 DB 1739 ATGATCGAATGTCCAGGAAAGGTTTTTGAAGCCAAATGCTTGTGAAATAGTAAATGAA 1798
 QY 1518 CAAATTTAAATGACACCGAGTATCTGATTTTCAAGAAAAAACAATTTGAATTTGGTTTCCC 1577
 DB 1799 CAAATTTAAATGACACCGAGTATCTGATTTTCAAGAAAAAACAATTTGAATTTGGTTTCCC 1858
 QY 1578 GAAAAAGAACTTTTGTGCTGCTGTTTTTGTGATGTAACGAAACAGCGGGGAAATCCATGCTTA 1637

[illegible]

Db	2936	ATGGAGAAAGATCCAGAGCCGAAGTACACGTTTCATTGTGATTCAGAAAGAGACAATATAC	2995
Qy	2572	GATTCCTTCGAAGAAATGAAAGATAGCCAGTGGTCAATAAAGATCTTACTCTCGCTG	2631
Db	2996	GATTCCTTCGAAGAAATGAAAGATAGCCAGT-GTCAATAAAGATCTTACTCTCGCTG	3054
Qy	2632	AAACAGATGTCGCTGTTGCTCTGTATAACAATGGGAGGAGATATGAAAGAAAGCAAAG	2691
Db	3055	AAACAGATGTCGCTGTTGCTCTGTATAACAATGGGAGGAGATATGAAAGAAAGCAAAG	3114
Qy	2692	AAACTGGAATTTGTGAACCCATCTCCGGAAACAATGTGATATAACTTATCGTTTCGAAAT	2751
Db	3115	AAACTGGAATTTGTGAACCCATCTCCGGAAACAATGTGATATAACTTATCGTTTCGAAAT	3174
Qy	2752	ACAAATTCGATTTTTTCTTGGCATCTCATCATGTGTGTCCTTGGTACATCTCGTCCAGGAC	2811
Db	3175	ACAAATTCGATTTTTTCTTGGCATCTCATCATGTGTGTCCTTGGTACATCTCGTCCAGGAC	3234
Qy	2812	ATTACACTGTTATGTATGACGATAAAGGAATGAGCCAAGATGAAGTCTAT-----	2861
Db	3235	ATTACACTGTTATGTATGACGATAAAGGAATGAGCCAAGATGAAGTCTATGTAAGCGTTT	3294
Qy	2862	-----	2861
Db	3295	TGAATPAGCAGTTAGCGATTTTATGATTTTGTAAATCCGCATATAGTTATTATTAATAAAAAATG	3354
Qy	2862	-----AAATGACCTACGGACTGTCCTTCTCTCTGCTAGATGTCGAAACCCATCTCG	2915
Db	3355	TTTCAGAAATGACCTACGGACTTGTCTCTCTCTAGATGTCGAAACCCATCTCG	3414
Qy	2916	TGCTCTGTTCCGGTTCAATTATGCTCATTTTATCATGTGAAAAAGCGAAAGCTTTTATCGA	2975
Db	3415	TGCTCTGTTCCGGTTCAATTATGCTCATTTTATCATGTGAAAAAGCGAAAGCTTTTATCGA	3474
Qy	2976	ACTTACAAGGAACAATTACATCGGTGACTATGCAAGCCACGGACTCGACAACGAAATGGAA	3035
Db	3475	ACTTACAAGGAACAATTACATCGGTGACTATGCAAGCCACGGACTCGACAACGAAATGGAA	3534
Qy	3036	CATTTTCTCCAAACTTAACGTTGAGTACCCCTGGAAATGTCGTCGCATTAACATTTTGCAAA	3095
Db	3535	CATTTTCTCCAAACTTAACGTTGAGTACCCCTGGAAATGTCGTCGCATTAACATTTTGCAAA	3594
Qy	3096	GTGTCGCCCGTTTCAATCAAAATTTTTCAATTTGTAGATATTGTACTTACTTTTTTTTAAAG	3155
Db	3595	GTGTCGCCCGTTTCAATCAAAATTTTTCAATTTGTAGATATTGTACTTACTTTTTTTTAAAG	3654
Qy	3156	CCCGGTTTCAAAATTCATTTCCATGACTAACGGTTTTTCATAAATTAATTGAATTT	3210
Db	3655	CCCGGTTTCAAAATTCATTTCCATGACTAACGGTTTTTCATAAATTAATTGAATTT	3709

RESULT 3

RESULI 3
ADB17459

ID ADB17459 standard; cDNA; 2827 BP.

[illegible]

AC ADB17459;

XX
1500/1500

DT 20-NOV-2003 (first entry)

XX

DE Soybean post-transcriptional gene silencing Zwiille homolog cDNA #1.

XX

KW Soybean; ss; gene; post-transcriptional gene silencing; PTGS;

KW RNA degradation; transgene expression; plant; transgenic plant;

KW transposon activity; meristem activity; architecture; development;

KW proliferation; callus propagation.

338

OS Glycine max.

THE

FH	Key	Location
FTH	CHC	2 3467

FT	CDS	2.	.2467
FT		14439	

```

EE / (a34477)
EE /*tag=3

```

13

QY 2674 ATATGAAGAAGCAAGAACTGGAATTTGTGAACCCCATCTCCGGAACAACCTGTGGATA 2733
 Db 2028 ACAGGAGCAGAGCAGTACAGATCGAGTGGGATATATTCCTGGGACTGTGTGTGATA 2087
 QY 2734 AACTTATCGTTTCGAAATACAAATTCGAATTTTCTTGGCATCTCATCATGTGTGTCCTTG 2793
 Db 2088 CCAAAATCTGCCATCCACAGAAATTTGATTTTATCTCTGCAGCCATCTGCGATCCAGG 2147
 QY 2794 GTACATCTCGTCCAGGACATACACTGTATGTATGATGAGTAAAGGAATGAGCCAGATG 2853
 Db 2148 GTACTAGTCGCGCAGCTCATTTATCATGTGCTGTGGGATGAAACAACTTTCACCTGATG 2207
 QY 2854 AAGTCTATAAATGACCTACGACTTGTCTTTCTCTCTGTAGATGTGGAACCCATCT 2913
 Db 2208 GAATTCAGTCTCTGACAAACAACTTTGTTATACATATGCGAGGTGTACACGCTCAGTAT 2267
 QY 2914 CGTTGCCCTGTTCCGGTTCATTTATGCTCATTTATCA 2948
 Db 2268 CAGTTGTTCTCCAGCATATTATGCACATTTAGCA 2302

RESULT 5

ACH23411
 ID ACH23411 standard; cDNA; 490 BP.

AC ACH23411;

XX 13-OCT-2003 (first entry)

DE Human adult ovary cDNA #1791.

KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

FI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 10623; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide

CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC sequence.uspto.gov/sequence.html?DocID=20030073623
 XX
 SQ Sequence 490 BP; 135 A; 107 C; 109 G; 135 T; 0 U; 4 Other;

Query Match 1.9%; Score 62.4; DB 9; Length 490;
 Best Local Similarity 51.8%; Pred. No. 0.00023;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATGTGAACCCATCATCCGGAACAACCTGTGGATAAACTATATCGTTTCGAAATACAAAT 2757
 Db 215 GAAGTGGCAATATCCAGCTGGAAACAACAGTTGTATACAGACATTTACACACCCATATGAT 274
 QY 2758 TCGATTTTCTTCTGCGCATCTCATCATGTGTGCTTGTGGTACATCTCGTCCAGGACATTACA 2817
 Db 275 TCGATTTTCTTCTGCGCATCTCATCATGTGTGCTTGTGGTACATCTCGTCCAGGACATTACA 334
 QY 2818 CTGTTATGTATGACGATAAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACGGAC 2877
 Db 335 ATGTTTATGGGATGATAAAGTCTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 394
 QY 2878 TTGCTTTTCTCTGCTAGATGTGCAAAACCCATCTCGTTCGCTGTTCCGGTTCATTATG 2937
 Db 395 TCTGCCACACTTACGTACGCTGTACAGCATCTGTTTCTATACCTGCACCGGATATTATG 454
 QY 2938 CTCATTTATCATGTGAAAGGCGAAGAGCTT 2969
 Db 455 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 486

RESULT 6

AAZ15059
 ID AAZ15059 standard; cDNA; 734 BP.

XX AAZ15059;

XX 12-OCT-1999 (first entry)

XX Human gene expression product cDNA sequence SEQ ID NO:2528.

XX Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

OS Homo sapiens.

XX WO9938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-US001619.

XX 28-JAN-1998; 98US-0072910P.

XX 24-FEB-1998; 98US-0075954P.

XX 31-MAR-1998; 98US-0080114P.

XX 03-APR-1998; 98US-0080515P.

XX 03-APR-1998; 98US-0080666P.

XX 21-OCT-1998; 98US-0105234P.

XX 28-OCT-1998; 98US-0105877P.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 XX WPI; 1999-494092/41.

XX

PT Novel human genes and their expression products which are differentially
PT expressed in different cell types.

PS Claim 1; Page 1234; 2479pp; English.

XX The present invention describes a library of human polynucleotides
XX comprising the sequences given in AAZ12532 to AAZ1779. Also described is
XX a method of detecting differentially expressed genes correlated with the
XX cancerous state of a mammalian cell, comprising detecting at least one
XX differentially expressed gene product in a test sample from a cell
XX suspected of being cancerous, where the gene product is encoded by one of
XX the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The
XX polynucleotides can be used as a source of primers and probes, which can
XX be used for a variety of purposes, e.g. detection of expression levels,
XX mapping, tissue typing or profiling, forensics, genetic analysis and
XX detection of polymorphisms. Polypeptides encoded by the polynucleotides
XX can be used for raising antibodies for experimental, diagnostic and
XX therapeutic purposes. The polynucleotides may also be used to construct
XX arrays for diagnostics (which may be used to determine function of an
XX encoded protein); and to detect differences in expression levels between
XX two cells (e.g. to identify abnormal or diseased tissue in a human, to
XX identify a genetic predisposition or susceptibility to a disease such as
XX cancer). The polynucleotides of the invention are especially used in the
XX diagnosis, prognosis and management of colorectal cancer, breast cancer,
XX and lung cancer. The polynucleotides can also be used to screen for
XX peptide analogues and antagonists

XX Sequence 734 BP; 219 A; 159 C; 155 G; 193 T; 0 U; 8 Other;

Query Match 1.9%; Score 62.4; DB 2; Length 734;

Best Local Similarity 51.8%; Pred. No. 0.00027;

Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757

DB 209 GAAGTGGCAATATCCAGCTGGAACAAACAGTTGATACAGACATTACACACCCATATGACT 268

QY 2758 TCGATTTTCTTGGCATCTCATCTGCTGGTACATCTCGTCCAGACATTACA 2817

DB 269 TCGATTTTACCTCTGTAGCCATGCTGGAATACAGGGTACCACTGCTCTTACACACTATC 328

QY 2818 CTGTTATGTATGACGATAAGGAATGAGCCAGATGAAGTCTATAAAATGACCTACGGAC 2877

DB 329 ATGTTTATGGATGATGATCTGTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 388

QY 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGCTGTTCCGGTTTCATTATG 2937

DB 389 TCTGCCACACTTACGCTACGCTGTACACGATCTGTTTCTATACCTGCACACGCGTATTATG 448

QY 2938 CTCATTATCATGTGAAAGCGAAGAGCTT 2969

DB 449 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 480

ADJ94706

ADJ94706;

RESULT 7

ADJ94706

ID ADJ94706 standard; cDNA; 2568 BP.

XX AC

XX ADJ94706;

XX 06-MAY-2004 (first entry)

XX DE

XX Human eukaryotic translation initiation factor 2C4 (eIF2C4) cDNA.

XX DE

XX cytotostatic; antiinflammatory; virucide; immunosuppressive; tumour;

XX inflammatory; infectious disease; viral infection; degenerative;

XX autoimmune; gene therapy; Argonaute family;

XX eukaryotic translation initiation factor 2C4; eIF2C4; human; ss; gene.

XX OS

XX Homo sapiens.

XX PN

XX WO2004007718-A2.

XX

PD 22-JAN-2004.

XX PF

XX 10-JUL-2003; 2003WO-EP007516.

XX PR

XX 10-JUL-2002; 2002EP-00015532.

XX PR

XX 23-AUG-2002; 2002EP-00018906.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PA

XX Tuschl T, Martinez J, Patkaniowska A, Urlaub H, Luehrmann R;

XX PI

XX WPI; 2004-122948/12.

XX DR

XX P-PSDB; ADJ94700.

XX

XX New single-stranded RNA molecule having a length from 14-50 nucleotides,

XX PT

XX useful for preventing or treating tumor, inflammatory, infectious, e.g.

XX PT

XX viral infections, degenerative and autoimmune diseases.

XX

XX Example; Fig 16; 73pp; English.

XX

XX The invention relates to a novel single-stranded RNA molecule having a

XX length from 14-50 nucleotides where at least 14-20 of the 5'-most

XX nucleotides are substantially complementary to a target transcript. The

XX RNA molecule of the invention demonstrates cytostatic, antiinflammatory,

XX virucide and immunosuppressive activities and may be useful for

XX inhibiting the expression of a target gene in vitro or in vivo,

XX preferably for preventing or treating diseases associated with the

XX overexpression of at least one target transcript. The diseases may be

XX selected from tumour diseases, inflammatory diseases and autoimmune diseases.

XX such as viral infections, degenerative diseases and autoimmune diseases.

XX Furthermore, the molecules of the invention may be utilised during gene

XX therapy. The current sequence is that of the human eukaryotic translation

XX initiation factor 2C4 (eIF2C4) cDNA of the invention.

XX

XX SQ

XX Sequence 2568 BP; 753 A; 568 C; 600 G; 647 T; 0 U; 0 Other;

Query Match 1.9%; Score 62.4; DB 12; Length 2568;

Best Local Similarity 51.8%; Pred. No. 0.00042;

Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757

DB 2165 GAAGTGGCAATATCCAGCTGGAACAAACAGTTGATACAGACATTACACACCCATATGACT 2224

QY 2758 TCGATTTTCTTGGCATCTCATCTGCTGGTACATCTCGTCCAGGACATTACA 2817

DB 2225 TCGATTTTACCTCTGTAGCCATGCTGGAATACAGGGTACCACTGCTCTTACACTATC 2284

QY 2818 CTGTTATGTATGACGATAAGGAATGAGCCAGATGAAGTCTATAAAATGACCTACGGAC 2877

DB 2285 ATGTTTATGGATGATGATGCTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 2344

QY 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGCTGTTCCGGTTTCATTATG 2937

DB 2345 TCTGCCACACTTACGCTACGCTGTACACGATCTGTTTCTATACCTGCACACGCGTATTATG 2404

QY 2938 CTCATTATCATGTGAAAGCGAAGAGCTT 2969

DB 2405 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2436

RESULT 8

AAK53425

ID AAK53425 standard; cDNA; 2914 BP.

XX AC

XX AAK53425;

XX 06-NOV-2001 (first entry)

XX

XX Human polynucleotide SEQ ID NO 3938.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX

KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.

XX WO200157190-A2.
 XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.
 XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.
 XX 20-JUN-2000; 2000US-00598075.
 XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.
 XX 15-SEP-2000; 2000US-00663561.
 XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.
 XX (HYSB-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX

DR WPI; 2001-476283/51.
 DR P-PSDB; AAM80292.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.

XX Claim 1; Page 6185-6188; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX

SQ Sequence 2914 BP; 856 A; 638 C; 668 G; 752 T; 0 U; 0 Other;

Query Match 1.9%; Score 62.4; DB 4; Length 2914;
 Best Local Similarity 51.8%; Pred. No. 0.00044;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATGTCGACCATTCGCGAACACTGTCGTAACCTTATCGTTTGAATACAAAT 2757
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2180 GAATGGCATATCCAGCTGGGAACCACTGTTGATACAGACATTACACCCCATATGAGT 2239
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2758 TCGATTTTCTTGGCATCTCATCATGCTGCTGTCATATCTCGTCCAGGACATTACA 2817
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2240 TCGATTTTACCTCTGTAGCCATGCTGGAATACAGGGTACCATGCTCTCACACTATC 2299
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2818 CTGTTATGATGACGATAAGGAATAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2300 ATGTTTATGGGATGATACTGCTTTACTGCAGATGAACCTTACGCTGAACCTTACCAGC 2359
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2878 TTGCTTTTCTCTGCTAGATGTCGAACCCCATCTGTTGCTGTTCCGGTTCATTATG 2937
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2360 TTGCCACACTTACGTACGCTGTACAGCATCTGTTTCTATACCTGCACCGCGTATTATG 2419
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2938 CTCATTATATCATGTGAAAGCGAAAGAGCTT 2969
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2420 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2451

RESULT 9
 ABX13625

ID ABX13625 standard; cDNA; 2914 BP.

XX AC ABX13625;

XX 12-FEB-2003 (first entry)

XX Human cytokine or cell proliferation/differentiation protein cDNA #2.

XX Human; ss; gene; cytokine; cell proliferation; cell differentiation;
 KW anaemia; wound; ulcer; thrombocytopenia; osteoporosis; osteoarthritis;
 KW inflammation; Alzheimer's disease; Parkinson's disease; stroke;
 KW Huntington's disease; amyotrophic lateral sclerosis; HIV;
 KW immune deficiency; human immunodeficiency virus infection;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW rheumatoid arthritis; Guillain-Barre syndrome; graft-versus-host disease;
 KW cancer; thyroid cancer; lung cancer; small cell carcinoma;
 KW Kaposi's sarcoma; brain tumour; prostate cancer; ovarian cancer;
 KW leukaemia; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; food supplement.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..2583

XX FT /*tag= a

XX FT /product= "Cytokine/proliferation/differentiation protein

XX FT #2"

XX US2002137044-A1.

XX 26-SEP-2002.

XX 30-JAN-2001; 2001US-00774434.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX (TANG/) TANG Y T.

XX (ZHOU/) ZHOU P.

XX (GOOD/) GOODRICH R.

XX (LIUC/) LIU C.

XX (ASUN/) ASUNDI V.

XX (WEHR/) WEHRMAN T.

XX (YANG/) YANG Y.

XX (DRMA/) DRMANAC R T.

XX Tang YT, Zhou P, Goodrich R, Liu C, Asundi V, Wehrman T, Yang Y;
 PI Drmanac RT;
 XX WPI; 2003-110596/10.
 XX P-PSDB; ABG72612.

XX New polynucleotides for diagnosing, treating or preventing e.g. anemia,
 PT wounds, ulcers, thrombocytopenia, osteoporosis, inflammations,
 PT Alzheimer's disease, stroke, autoimmune disorders or cancers.

XX Claim 1; Page 46-49; 63pp; English.

XX The invention relates to an isolated polynucleotide encoding a cytokine
 CC or cell proliferation/differentiation-related protein (or the mature
 CC protein or active domain), sequences greater than 90% identical to or
 CC sequences complementary to it. Also included are the encoded proteins,
 CC expression vectors, transformed host cells, antibodies, and identifying a
 CC compound that binds to the polypeptide. The polynucleotide, polypeptides
 CC encoded by it, or antibodies to the polypeptides are useful in
 CC therapeutic, diagnostic or research methods. They are particularly useful
 CC for diagnosing, treating or preventing e.g. anaemias, wounds, ulcers,
 CC thrombocytopenia, osteoporosis, osteoarthritis, inflammations,
 CC Alzheimer's disease, Parkinson's disease, Huntington's disease,

CC amytrophic lateral sclerosis, stroke, immune deficiencies (e.g. human
CC immunodeficiency virus (HIV), severe combined immunodeficiency or
CC infections), autoimmune disorders (e.g. rheumatoid arthritis, Guillain-
CC Barre syndrome or graft-versus-host disease), cancers (e.g. thyroid
CC cancer, lung cancers, small cell carcinoma, Kaposi's sarcoma, brain
CC tumours, prostate cancer, ovarian cancer or leukaemia), or inflammatory
CC disorders (e.g. inflammatory bowel disease or Crohn's disease). The
CC polynucleotides and proteins are useful for screening peptides or small
CC molecule inhibitors or agonists that are useful for treating these
CC diseases. The polypeptide is also useful as molecular markers, or as a
CC food supplement. The present sequence is one of the human polynucleotides
CC of the invention encoding a cytokine or cell proliferation/
CC differentiation-related protein

XX SQ Sequence 2914 BP; 856 A; 638 C; 668 G; 752 T; 0 U; 0 Other;

Query Match 1.9%; Score 62.4; DB 8; Length 2914;
Best Local Similarity 51.8%; Pred. No. 0.00044;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGTAACTTATCGTTTCGAAATACAAAT 2757
Db 2180 GAAGTGGCAATATCCAGCTGGAAACAAAGTTGATACAGACATTACACACCCCATATGAGT 2239
QY 2758 TCGATTTTCTTGGCATCTCATCATGTTGCTTGGTACATCTCGTCCAGGACATTACA 2817
Db 2240 TCGATTTTACCTCTGTAGCCATCTGCGAATACAGGGTACCACTCGTCTTCACACTATC 2299
QY 2818 CTGTTATGTATGACGATAAGGAATGAGCAAGTGAAGTCTATAAATGACCTACGGAC 2877
Db 2300 ATGTTTATGGATGATGATGCTTCTTACTGCAGATGAATCTCAGCTGCTAATACCAAGC 2359
QY 2878 TTGCTTTTCTCTGTAGATGTCGAAACCCATCTCGTTGCTGTTCGGTTTCATTATG 2937
Db 2360 TCTGCCACACTTACGTACGCTGTACACGATCTGTTTCTATACCTGCACCAAGGATATG 2419
QY 2938 CTCATTATCATGTGAAAGCGAAGAGCTT 2969
Db 2420 CTCACCTGTAGCATTTAGAGCCAGATATCAT 2451

RESULT 10
AAH16058
ID AAH16058 standard; cDNA; 3050 BP.

XX AC AAH16058;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:14738.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000BP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
XX Isehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.

XX PS Claim 8; SEQ ID NO 14738; 2537pp + Sequence Listing; English.

XX CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX SQ Sequence 3050 BP; 886 A; 687 C; 702 G; 775 T; 0 U; 0 Other;

Query Match 1.9%; Score 62.4; DB 4; Length 3050;
Best Local Similarity 51.8%; Pred. No. 0.00045;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCATCATCCGGAACAACTGTGGTAACTTATCGTTTCGAAATACAAAT 2757
Db 2280 GAAGTGGCAATATCCAGCTGGAAACAAAGTTGATACAGACATTACACACCCCATATGAGT 2339
QY 2758 TCGATTTTCTTGGCATCTCATCATGTTGCTGTTGGTACATCTCGTCCAGGACATTACA 2817
Db 2340 TCGATTTTACCTCTGTAGCCATCTGGAATACAGGTTACAGCTCGTCTTCACACTATC 2399
QY 2818 CTGTTATGTATGACGATAAAGAAATGAGCCAGATGAACTCTATAAATGACCTACGGAC 2877
Db 2400 ATGTTTATGGATGATGATGCTTCTTACTGCAGATGAACCTTCTGCTGCTAATACCAAGC 2459
QY 2878 TTGCTTTTCTCTCTAGATGTCGAAACCCATCTCGTTGCTGTTCCGTTTCATTATG 2937
Db 2460 TCTGCCACACTTACGTACGCTGTACAGATCTGTTTCTATACCTGCACCAAGCGTATTATG 2519
QY 2938 CTCATTATCATGTGAAAGCGAAGAGCTT 2969
Db 2520 CTCACCTGTAGCATTTAGAGCCAGATATCAT 2551

RESULT 11
ADH77114
ID ADH77114 standard; cDNA; 3050 BP.

XX AC ADH77114;
XX 22-APR-2004 (first entry)
XX Human PAZ/PIWI domain-containing protein cDNA.
XX cytostatic; PAZ/PIWI domain-containing protein inhibitor;
XX PAZ/PIWI domain-containing protein; hyperproliferative disorder; cancer;
XX aberrant cellular differentiation; human;
XX PAZ/PIWI domain-containing protein; gene; ss.
XX

OS Homo sapiens.
 PN US2003232442-A1.
 XX 18-DEC-2003.
 PD 17-JUN-2002; 2002US-00175492.
 PF 17-JUN-2002; 2002US-00175492.
 PR (ISIS-) ISIS PHARM INC.
 XX Dobie KW;
 XX WPI; 2004-052174/05.
 DR P-PSDB; ADH77275.
 XX New antisense oligonucleotide targeted to a nucleic acid encoding a
 PT PAZ/PIWI domain-containing protein, useful for treating cancer or a
 PT disease arising from aberrant cellular differentiation.
 XX Example 13; SEQ ID NO 4; 119pp; English.
 XX The invention describes a compound 8-80 nucleobases in length targeted
 CC to, and which specifically hybridizes with a nucleic acid molecule
 CC encoding a PAZ/PIWI domain-containing protein, and inhibits the
 CC expression of a PAZ/PIWI domain-containing protein. The compound,
 CC associated with PAZ/PIWI domain-containing protein, such as a
 CC hyperproliferative disorder e.g. cancer, or a disease or condition
 CC arising from aberrant cellular differentiation. They are also useful in
 CC research and diagnostics for modulating the expression of PAZ/PIWI domain
 CC -containing protein. This sequence encodes a human PAZ/PIWI domain-
 CC containing protein.
 XX
 SQ Sequence 3050 BP; 886 A; 687 C; 702 G; 775 T; 0 U; 0 Other;
 Query Match 1.9%; Score 62.4; DB 12; Length 3050;
 Best Local Similarity 51.8%; Pred. No. 0.00045;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 2698 GAATTGGAACCCATCATCCGGAACAACCTGGGATAAACTTATCGTTTCGAAATACAAAT 2757
 DB 2280 GAAGTGCATATCCGAGCTGGAAACAACAGTTGATACAGATTTACACACCCATATGAT 2339
 QY 2758 TCGATTTTCTTGGCATCTCATCATGGTGTCTTGGTATCATCTCGTCCAGGACATTACA 2817
 DB 2340 TCGATTTTACCTCTGTAGCCATCTCGAATACAGGCTACAGTCTGCTTCCACACTATC 2399
 QY 2818 CTGTTATGTATGACGATAAGGATGAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
 DB 2400 ATGTTTATGGATGATACTGCTTTACTCGAGATGAACCTTCAGCTGCTAACTTACCAGC 2459
 QY 2878 TTGCTTTTCTCTGTAGATGTGGAACCCATCTCGTTCCTGCTTCCGTTTCATTATG 2937
 DB 2460 TCTGCCACACTTACGTACGCTGTACAGATCTGTTTCTATACCTTGCACCCAGGTATTATG 2519
 QY 2938 CTCATTATCATGTGAAAGGCGAAGAGCTT 2969
 DB 2520 CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2551
 RESULT 12
 AAD55860
 ID AAD55860 standard; cDNA; 3325 BP.
 XX
 AC AAD55860;
 XX
 DT 07-AUG-2003 (first entry)
 XX Human nucleic acid associated protein (NAAP)-31 cDNA.
 DE Human; nucleic acid associated protein; NAAP; stroke; AIDS; nontropic;
 KW

KW cancer; atherosclerosis; neurological; epilepsy; Huntington's disease;
 KW developmental disorder; antiinflammatory; neuroprotective; thymomimetic;
 KW Cushing's syndrome; infection; gene therapy; cytostatic; anticonvulsant;
 KW cerebroprotective; hypothyroidism; cell proliferative disorder; allergy;
 KW gene; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 497..2947
 FT /*tag= a
 FT /product= "Human NAAP protein"
 XX
 PN WO2003006618-A2.
 XX
 PD 23-JAN-2003.
 XX
 PF 10-JUL-2002; 2002WO-US021971.
 XX
 PR 12-JUL-2001; 2001US-0305089P.
 PR 12-JUL-2001; 2001US-0305104P.
 PR 13-JUL-2001; 2001US-0305325P.
 PR 13-JUL-2001; 2001US-0305390P.
 PR 19-JUL-2001; 2001US-0306960P.
 PR 20-JUL-2001; 2001US-030694P.
 PR 27-JUL-2001; 2001US-0308170P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX Swarnakar A, Richardson TW, Warren BA, Griffin JA, Tang YT;
 PI Yue H, Baughn MR, Emerling BM, Lal PG, Lu DAM, Forsythe LJ;
 PI Ramkumar J, Li JX, Becha SD, Duggan BM, Sanjanwala MM, Lee EA;
 PI Burford N, Elliott VS, Ison CH, Ding L, Borowsky ML, Yao MG;
 PI Barroso I, Tran B, Walia NK, Hafalia AJA, Nguyen DB, Lu Y;
 PI Arvizu CS;
 XX
 DR WPI; 2003-221732/21.
 XX P-PSDB; AAE37044.
 XX
 PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
 PT infections.
 XX
 PS Claim 5; Page 255-256; 260pp; English.
 XX
 CC The invention relates to human nucleic acid associated proteins (NAAP)
 CC and their corresponding nucleic acid sequences. The invention is useful
 CC in diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of NAAP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. It is also useful in assessing the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of NAAP. The NAAP or its fragments are useful in screening
 CC compounds for which acts as their agonist or antagonist. The microarray
 CC is useful in monitoring or measuring protein-protein interactions, drug-
 CC target interactions, and gene expression profiles. NAAP DNA is used in
 CC gene therapy. The present sequence is human NAAP cDNA
 XX
 SQ Sequence 3325 BP; 902 A; 809 C; 813 G; 801 T; 0 U; 0 Other;
 Query Match 1.9%; Score 62.4; DB 8; Length 3325;
 Best Local Similarity 51.8%; Pred. No. 0.00046;
 Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
 QY 2698 GAATTGGAACCCATCATCCGGAACAACCTGGGATAAACTTATCGTTTCGAAATACAAAT 2757
 DB 2544 GAAGTGCATATCCGAGCTGGAAACAACAGTTGATACAGACATTACACACCCATATGAT 2603
 QY 2758 TCGATTTTCTTGGCATCTCATCATGGTGTCTTGGTATCATCTCGTCCAGGACATTACA 2817

Db 2604 TCGATTTTACCTCTGCTGACCATGCTGGATACAGGGTACCAGTCTGCTTACACTATC 2663
Qy 2818 CTGTTATGATGACGATAAGGAATGAGCAAGATGAATCTATAAAATGACCTACGGAC 2877
Db 2664 ATGTTTATGGGATGATACTGCTTTACTGCGAGATGAATCTTACGCTGCTAACTTACCAGC 2723
Qy 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTGGTCTGCTTCCGGTTTCAATTATG 2937
Db 2724 TCTGCCACACTTACGTACGCTGTCACAGATCTGTTTCTATACCTGCACCGCGTATATG 2783
Qy 2938 CTCATTATCATGTGAAAAAGCGAAAGAGCTT 2969
Db 2784 CTCACCTGGTAGCATTTAGACCGACAGATATCAT 2815

RESULT 13
ABX34752
ID ABX34752 standard; cDNA; 3736 BP.
XX AC ABX34752;
XX
DT 13-FEB-2003 (first entry)
XX
DE Human mddt cDNA SEQ ID 313.
XX
KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200279449-A2.
XX
PD 10-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-US009944.
XX
PR 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
(INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JR, Yu JY, Tuason O, Yap PE, Anshey SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DH, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
XX WPI: 2003-058431/05.
XX P-PSDB; ABU11762.
XX
XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT [e.g. leukemia or sarcoma], anaemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis.
XX
PS Claim 1; SEQ ID NO 313; 339pp + Sequence Listing; English.
XX
XX This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
CC

CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy, of
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABX34440-ABX34835 encode the MDDT polypeptides represented in
CC ABU11450-ABU11845, described in the disclosure of the invention. NOTE:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 3736 BP; 1061 A; 853 C; 809 G; 1013 T; 0 U; 0 Other;
Query Match 1.9%; Score 62.4; DB 8; Length 3736;
Best Local Similarity 51.8%; Pred. No. 0.00048;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
Qy 2698 GAATTGTGAACCCATCATCCGGAACAACACTGTGGATAAACTTATCGTTTGGAAATACAAAT 2757
Db 2445 GAAGTGGCAATATCCAGCTGGGAACAACAGTTTGATACAGACATTACACACCCCATATGAGT 2504
Qy 2758 TCGATTTTCTTGGCACTCATCATGCTGCTTGGTACATCTCGTCCAGGACATTACA 2817
Db 2505 TCGATTTTACCTCTGTAGCCATGTGGAAATACAGGTTACCGTCTCTTACACTATC 2564
Qy 2818 CTGTTATGTATCAGCATAAAGGAATGAGCAAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 2565 ATGTTTATGGATGATATACTGCTTTACTGCGAGTGAATCTTACGCTGCTTAATTACAGC 2624
Qy 2878 TTGCTTTTCTCTGCTAGATTCGAAACCCATCTCGTTGCTGTTCGGTTCATATG 2937
Db 2625 TCTGCCACACTTACGTACGCTGTACAGATCTGTTTCTATACCTGCACCGCGTATTATG 2684
Qy 2938 CTCATTATCATGTGAAAAGCGAAAGAGCTT 2969
Db 2685 CTCACCTGGTAGCATTTTAGAGCCAGATATCAT 2716

RESULT 14
ADQ80688
ID ADQ80688 standard; DNA; 2679 BP.
XX AC ADQ80688;
XX
DT 21-OCT-2004 (first entry)
XX
DE Arabidopsis thaliana TFL1-binding protein coding sequence #2.
XX
KW TFL1-binding protein; plant growth control; biotechnology;
KW fishing industry; screening; gene; ds.
XX Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
FH CDS 1..2679
FT /*tag= a
FT /product= "Arabidopsis thaliana TFL1-binding protein #2"
XX
PN JP2004208572-A.
XX
PD 29-JUL-2004.
XX
XX 27-DEC-2002; 2002JP-00381220.
XX
XX 27-DEC-2002; 2002JP-00381220.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX WPI; 2004-538191/52.
DR

Db 1091 CATTGGCAGCTTACAGAGC 1110

Search completed: July 5, 2006, 21:52:46
Job time : 1730 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 22:08:48 ; Search time 3761 Seconds
(without alignments)
10542.994 Million cell

```

Title: US-10-645-746-2
Perfect score: 3227
Sequence: 1 cagccacaagtgtgaac.....tttaaaaaaaaaaaaaaaaa 3227

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA_Main:*

- 1: /EMC Celleria SIDS3/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC Celleria SIDS3/prodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC Celleria SIDS3/prodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC Celleria SIDS3/prodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC Celleria SIDS3/prodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC Celleria SIDS3/prodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC Celleria SIDS3/prodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC Celleria SIDS3/prodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC Celleria SIDS3/prodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC Celleria SIDS3/prodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC Celleria SIDS3/prodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC Celleria SIDS3/prodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC Celleria SIDS3/prodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC Celleria SIDS3/prodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC Celleria SIDS3/prodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC Celleria SIDS3/prodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			Description		
	Score	Match	Length	DB	ID	Description
1	3227	100.0	3227	9	US-10-645-745-2	Sequence 2, Appli
2	3227	100.0	3227	10	US-10-645-735-2	Sequence 2, Appli
3	3227	100.0	3227	16	US-11-144-985-2	Sequence 2, Appli
4	2601	80.6	3719	9	US-10-645-746-1	Sequence 1, Appli
5	2601	80.6	3719	10	US-10-645-735-1	Sequence 1, Appli
6	2601	80.6	3719	16	US-11-144-985-1	Sequence 1, Appli
7	63.8	2.0	2827	6	US-10-174-363-3	Sequence 3, Appli
8	63.8	2.0	2827	13	US-11-093-888-3	Sequence 3, Appli
9	62.4	1.9	490	3	US-09-918-995-10623	Sequence 10623, A
10	62.4	1.9	734	10	US-10-779-543-3372	Sequence 3372, A
11	62.4	1.9	490	3	US-09-918-995-10623	Sequence 2, Appli
12	62.4	1.9	2914	3	US-09-774-434-2	Sequence 2, Appli
13	62.4	1.9	3050	7	US-10-175-493-4	Sequence 4, Appli
14	62.4	1.9	3325	10	US-10-483-505-66	Sequence 66, Appli
15	60	1.9	3578	10	US-10-756-149-37	Sequence 37, Appli
16	60	1.9	1501	6	US-10-174-363-5	Sequence 5, Appli
17	58.4	1.8	970	8	US-10-424-599-43001	Sequence 43001, A

ALIGNMENTS

RESULT 1

```

US-10-645-746-2
; Sequence 2, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATH
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMW-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21) ... (3080)
US-10-645-746-2

```

Query Match 100.0%; Score 3227; DB 9; Length 3227;

Best Local Similarity 100.0%;
 Matches 3227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 1 CAGCCACAAAGTGATGAAACATGTCCTCGAATTTCCTCCGAAATTCGGAAAAAGGATTTTATC 60

7

Db 1 CAGCCACAAAGTGATGAAACATGTCCTCGAATTTTCCGAATTGGAAAAGGATTTTATC 60

QV 61 GTCAATTCTTCGATCCGGAGATGAATGGCTTGCAGGGCCCACTGGTAATCGCAGCGCA 12

Qy	2281	AAGAAATGTCGTCCCGGTGAGCGTGCAGTGGCTCATGTGACGGGAAGAACAGATATTTTGG	2340
Db	2281	AAGAAATGTCGTCCCGGTGAGCGTGCAGTGGCTCATGTGACGGGAAGAACAGATATTTTGG	2340
Qy	2341	AAGCAAAGTTGCTGGAATTTGCTCAGAGAAATTCGACGAAAAACACGACAAATCGACACCAG	2400
Db	2341	AAGCAAAGTTGCTGGAATTTGCTCAGAGAAATTCGACGAAAAACACGACAAATCGACACCAG	2400
Qy	2401	CGCATATTGTAGTCTATCGAGACGGAGTTAGCGATTCCGGAGATTCGAGATGCTACGTTAGTCATG	2460
Db	2401	CGCATATTGTAGTCTATCGAGACGGAGTTAGCGATTCCGGAGATTCGAGATGCTACGTTAGTCATG	2460
Qy	2461	ATGAGCTTCGATCTTTAAAAAGCGAAGTAAAAACAATTCATGTCGGAAACGGGATCGGAGAAG	2520
Db	2461	ATGAGCTTCGATCTTTAAAAAGCGAAGTAAAAACAATTCATGTCGGAAACGGGATCGGAGAAG	2520
Qy	2521	ATCCAGACCGAAGTACACGTTTCATTTGTGATTCAGAAAAAGACACAATACACGATTTGCTTC	2580
Db	2521	ATCCAGACCGAAGTACACGTTTCATTTGTGATTCAGAAAAAGACACAATACACGATTTGCTTC	2580
Qy	2581	GAAGAAATCGAAAAAGATGAAGCCAGTGGTCAATAAAGATCTTACTCCCTGCTGAAACAGATG	2640
Db	2581	GAAGAAATCGAAAAAGATGAAGCCAGTGGTCAATAAAGATCTTACTCCCTGCTGAAACAGATG	2640
Qy	2641	TCGCTGTTGCTGCTGTTAAAAACAATGGGAGGAGGATATGAAAGAAACAAAGAAACCTGCAA	2700
Db	2641	TCGCTGTTGCTGCTGTTAAAAACAATGGGAGGAGGATATGAAAGAAACAAAGAAACCTGCAA	2700
Qy	2701	TTGTGAACCCATCATCCGGAAACAATGTGAGTAAACTTATCGTTTCGAAATACAAATTCG	2760
Db	2701	TTGTGAACCCATCATCCGGAAACAATGTGAGTAAACTTATCGTTTCGAAATACAAATTCG	2760
Qy	2761	ATTTTTCTTTGGCATCTCATCATGGTGTCTTGTGTAGATCTCGTCCAGGACATTTACACTG	2820
Db	2761	ATTTTTCTTTGGCATCTCATCATGGTGTCTTGTGTAGATCTCGTCCAGGACATTTACACTG	2820
Qy	2821	TTATGTATGCGATPAAAGGAATGAGCCAAGATGAAGTCTATAAATGACCTACCGGACTTG	2880
Db	2821	TTATGTATGCGATPAAAGGAATGAGCCAAGATGAAGTCTATAAATGACCTACCGGACTTG	2880
Qy	2881	CTTTTTCTCTGCTAGATGTCGAAAAACCCATCTCGTTGCCCTGTTCCGGTTCAATTATGCTC	2940
Db	2881	CTTTTTCTCTGCTAGATGTCGAAAAACCCATCTCGTTGCCCTGTTCCGGTTCAATTATGCTC	2940
Qy	2941	ATTTTATCATGTGAAAAACGGAAGAGCTTTTATCGAACTTACAGAGGAACAATTACATCGGTG	3000
Db	2941	ATTTTATCATGTGAAAAACGGAAGAGCTTTTATCGAACTTACAGAGGAACAATTACATCGGTG	3000
Qy	3001	ACTATGCAAGCCACGACTCGACACGAATTGGAAACATTTTCTCCAACTTAACGTAAGT	3060
Db	3001	ACTATGCAAGCCACGACTCGACACGAATTGGAAACATTTTCTCCAACTTAACGTAAGT	3060
Qy	3061	ACCCCTGGAATGTCGTTCCGATATACATTTTTCGAAAAAGTGCCTCCGGTTTCAATCAAAATTTT	3120
Db	3061	ACCCCTGGAATGTCGTTCCGATATACATTTTTCGAAAAAGTGCCTCCGGTTTCAATCAAAATTTT	3120
Qy	3121	TCAAATGTAGATATTGTACTTACTTTTTTTTAAAGCCCGGTTTCAAAAAATTCATCCATG	3180
Db	3121	TCAAATGTAGATATTGTACTTACTTTTTTTTAAAGCCCGGTTTCAAAAAATTCATCCATG	3180
Qy	3181	ACTAACGTTTTTCATAAAATTAATTGAAATTTAAAAAATAAAAAA	3227
Db	3181	ACTAACGTTTTTCATAAAATTAATTGAAATTTAAAAAATAAAAAA	3227

RESULT 2

RESULT 2
US-10-645-735-2

; Sequence 2, Application US/10645735

; Publication No. US20050100913A1

; GENERAL INFORMATION:

; APPLICANT: Mello, Craig C.

; APPLICANT: Tabara, Hiroaki

```

1  APPLICANT: Grishok, Alla
2  APPLICANT: Fire, Andrew
3  TITLE OF INVENTION: RNA INTERFERENCE PATH
4  TITLE OF INVENTION: INTERFERENCE
5  FILE REFERENCE: UMG-052
6  CURRENT APPLICATION NUMBER: US/10/645,735
7  CURRENT FILING DATE: 2003-08-20
8  PRIOR APPLICATION NUMBER: US/09/689,992A
9  PRIOR FILING DATE: 2000-10-13
10 PRIOR APPLICATION NUMBER: US 60/193,218
11 PRIOR FILING DATE: 2000-03-30
12 PRIOR APPLICATION NUMBER: US 60/159,776
13 PRIOR FILING DATE: 1999-10-15
14 NUMBER OF SEQ ID NOS: 14
15 SOFTWARE: FastSeq for Windows Version 4.0
16 SEQ ID NO 2
17 LENGTH: 3227
18 TYPE: DNA
19 ORGANISM: Caenorhabditis elegans
20 FEATURE:
21 NAME/KEY: CDS
22 LOCATION: (21)...(3080)
23 US-10-645-735-2

```

Query Match **100.0%**; **Score 3227**; **DB 10**; **Length 3227**;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	Qy	1	CAGCCCAAAAGTGATGAACAATGTCCTCGAATTTTCCGGAATTGGAAAAGGATTTTATC	60
	Db	1	CAGCCCAAAAGTGATGAACAATGTCCTCGAATTTTCCGGAATTGGAAAAGGATTTTATC	60
	Qy	61	GTCATTCTCTCGATCCGAGATGAAATGGCTTGCAGGCCCACTGGTAATGCGACGGCA	120
	Db	61	GTCATTCTCTCGATCCGAGATGAAATGGCTTGCAGGCCCACTGGTAATGCGACGGCA	120
	Qy	121	AATTCTATGAGAAGAAAGTACTTCTCTTTTGGTAAATTTGGTTCAAAGTTCTCCAGCAAAATTT	180
	Db	121	AATTCTATGAGAAGAAAGTACTTCTCTTTTGGTAAATTTGGTTCAAAGTTCTCCAGCAAAATTT	180
	Qy	181	ACGATCGGGAATACTACGAGTAGCAAGTGAAAAATGACAAAGGAAGTAGTATGAAATAGANAAC	240
	Db	181	ACGATCGGGAATACTACGAGTAGCAAGTGAAAAATGACAAAGGAAGTAGTATGAAATAGANAAC	240
	Qy	241	CAGGAAAAACCTTTCCCAAAAAAGACAGAAAATCCAAATTC CGGATCGTGCAAAAATCTTTCT	300
	Db	241	CAGGAAAAACCTTTCCCAAAAAAGACAGAAAATCCAAATTC CGGATCGTGCAAAAATCTTTCT	300
	Qy	301	GGCAACATCTTCGGCATGAGAAGAAGCAGACAGATTTTATTTCTCGAAGACTATGTTTTTG	360
	Db	301	GGCAACATCTTCGGCATGAGAAGAAGCAGACAGATTTTATTTCTCGAAGACTATGTTTTTG	360
	Qy	361	ATGAAAAAGGACACTGTTTATAGTGTGTCGCACTGAAACACTGTGCACATCAAAAATGCTGG	420
	Db	361	ATGAAAAAGGACACTGTTTATAGTGTGTCGCACTGAAACACTGTGCACATCAAAAATGCTGG	420
	Qy	421	TTTTCGGAGAAAGTAGTAAAAAAGGATTCGGAGAAAAAAGATGAAAAGGATTTGCGAGAAA	480
	Db	421	TTTTCGGAGAAAGTAGTAAAAAAGGATTCGGAGAAAAAAGATGAAAAGGATTTGCGAGAAA	480
	Qy	481	AAATCTTATACAAATGATACTTACCCTATCTGTAABAAATTTCCACCTCAACTTTTAGTCGAG	540
	Db	481	AAATCTTATACAAATGATACTTACCCTATCTGTAABAAATTTCCACCTCAACTTTTAGTCGAG	540
	Qy	541	AAAAATCCGGAAAAAGAACGAAGCGAATCCGAGTTACAAATTCCTGAAGAATGTTTATGA	600
	Db	541	AAAAATCCGGAAAAAGAACGAAGCGAATCCGAGTTACAAATTCCTGAAGAATGTTTATGA	600
	Qy	601	CCCAGAAAGTTGCGTACCGCGCTTTTGTGAAACGAGGAGATTAAAGTACAAATTCGCGAAA	660
	Db	601	CCCAGAAAGTTGCGTACCGCGCTTTTGTGAAACGAGGAGATTAAAGTACAAATTCGCGAAA	660
	Qy	661	ATTTTGTGTGACGATAATAAATCAATTTCTGCGAGTTCTCGAAATCGTTTTCACGATCCAAACA	720

[illegible]


```
QY 2881 CTTTCTCTCTCTAGATGTCGAAACCCCATCTCGTTGCCCTGTTCCGGTTCAATTATGCTC 2940
Db 2881 CTTTCTCTCTCTAGATGTCGAAACCCCATCTCGTTGCCCTGTTCCGGTTCAATTATGCTC 2940
QY 2941 ATTATCATGTGAAAGGCGAAGAGCTTTATCGAACTTCAAGGAACATTTACATCGGTG 3000
Db 2941 ATTATCATGTGAAAGGCGAAGAGCTTTATCGAACTTCAAGGAACATTTACATCGGTG 3000
QY 3001 ACTATGACAGCCACGAGCTCGACAGAAATGGAACATTTTCTCCTCAAACTAAACGTGAAGT 3060
Db 3001 ACTATGACAGCCACGAGCTCGACAGAAATGGAACATTTTCTCCTCAAACTAAACGTGAAGT 3060
QY 3061 ACCCTGGAATGCTGCGATACATTTTTCGAAAGTGTGCGCGTTTCAATCAAAATTTT 3120
Db 3061 ACCCTGGAATGCTGCGATACATTTTTCGAAAGTGTGCGCGTTTCAATCAAAATTTT 3120
QY 3121 TCAATGTAGATATTGTACTTACTTTTAAAGCCCGGTTTCAAAATTCATTCCTCATG 3180
Db 3121 TCAATGTAGATATTGTACTTACTTTTAAAGCCCGGTTTCAAAATTCATTCCTCATG 3180
QY 3181 ACTAAGCTTTTCAATAATCTGAAATTTTAAATAAAAAAAAAAAAAA 3227
Db 3181 ACTAAGCTTTTCAATAATCTGAAATTTTAAATAAAAAAAAAAAAAA 3227
```

RESULT 3

```
US-11-144-985-2
; Sequence 2, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: INTERFERENCE
; CURRENT APPLICATION NUMBER: US/11/144,985
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)....(3080)
US-11-144-985-2
```

```
Query Match 100.0%; Score 3227; DB 16; Length 3227;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCCAAAAGTGATGAACATGCTCTCGAATTTTCCGAAATGGAAAAAGGATTTTATC 60
Db 1 CAGCCCAAAAGTGATGAACATGCTCTCGAATTTTCCGAAATGGAAAAAGGATTTTATC 60
QY 61 GTCAATCTCTCGATCCGAGATGAAATGCTTCCGAGGCCCACTGCTTAATCGACGGCA 120
Db 61 GTCAATCTCTCGATCCGAGATGAAATGCTTCCGAGGCCCACTGCTTAATCGACGGCA 120
QY 121 AATCTATGAGAAGAAAGTACTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTT 180
Db 121 AATCTATGAGAAGAAAGTACTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTT 180
```

```
QY 181 ACGATCGGGAATACTACGAGTATGAAGTAAATGACAAAGGAAGTATTGAATAGAAAAC 240
Db 181 ACGATCGGGAATACTACGAGTATGAAGTAAATGACAAAGGAAGTATTGAATAGAAAAC 240
QY 241 CAGGAAACCTTTCCCAAAAAGACAGAAATTCCAATTCGCCGATCGTGCAAACTCTTCT 300
Db 241 CAGGAAACCTTTCCCAAAAAGACAGAAATTCCAATTCGCCGATCGTGCAAACTCTTCT 300
QY 301 GGCAACATCTTTCGCGATGAGAAGACAGACAGATTTTATTTCTCGAAGACTATGTTTTG 360
Db 301 GGCAACATCTTTCGCGATGAGAAGACAGACAGATTTTATTTCTCGAAGACTATGTTTTG 360
QY 361 ATGAAAAGGACACTGTTTATAGTGTTCGACGTGAACACATGTGCACATCAAAAATGCTGG 420
Db 361 ATGAAAAGGACACTGTTTATAGTGTTCGACGTGAACACATGTGCACATCAAAAATGCTGG 420
QY 421 TTTTCGAGAAAGTAGTAAATAAGGATTTCCGAGAAAAGAGATGAAAGGATTTGAGAAAA 480
Db 421 TTTTCGAGAAAGTAGTAAATAAGGATTTCCGAGAAAAGAGATGAAAGGATTTGAGAAAA 480
QY 481 AAATCTTTATACAAATGATATTTACCTTATCGTAAATAATTTACCTGAACTTTAGTCCAG 540
Db 481 AAATCTTTATACAAATGATATTTACCTTATCGTAAATAATTTACCTGAACTTTAGTCCAG 540
QY 541 AAATCCGAAAAGACGAAAGCGAATCGGAGTTTACAAATTCCTGAAGATGTTATGA 600
Db 541 AAATCCGAAAAGACGAAAGCGAATCGGAGTTTACAAATTCCTGAAGATGTTATGA 600
QY 601 CCCAGAAAGTTCCGTACGCGCTTTTGTGAACAGGAGATTTAAGTACAAATTCGCGAAA 660
Db 601 CCCAGAAAGTTCCGTACGCGCTTTTGTGAACAGGAGATTTAAGTACAAATTCGCGAAA 660
QY 661 ATTTGTGTACGATAAATTTCAATTTCTGCGAGTCTCTGAATCGTTTTCAGATCCAAACA 720
Db 661 ATTTGTGTACGATAAATTTCAATTTCTGCGAGTCTCTGAATCGTTTTCAGATCCAAACA 720
QY 721 GATTGCAACATCATTTAGAGTAGACCAAGAAATCGAAGCATGGTTTGGAAATTTACATG 780
Db 721 GATTGCAACATCATTTAGAGTAGACCAAGAAATCGAAGCATGGTTTGGAAATTTACATG 780
QY 781 GAATCAAAAGATTTGTTCCGATGGTGAACCTGTGCTCAATTTTGCATTTGCGATAAATCT 840
Db 781 GAATCAAAAGATTTGTTCCGATGGTGAACCTGTGCTCAATTTTGCATTTGCGATAAATCT 840
QY 841 TCTACAATGCACCGAAAATGTCTCTTCTGGAATTTCTCTCTTAATTTGTCGACCCCACT 900
Db 841 TCTACAATGCACCGAAAATGTCTCTTCTGGAATTTCTCTCTTAATTTGTCGACCCCACT 900
QY 901 CGTGTAAACGATGATGACGAAAGATCTTAAACAAACCTGATGCGCGGAAAAATGACAA 960
Db 901 CGTGTAAACGATGATGACGAAAGATCTTAAACAAACCTGATGCGCGGAAAAATGACAA 960
QY 961 TCAGACAACGCGCGCGCCCAAGATTTTCGACAAATTTTGGAAAATTTGAGCTGAAATGCG 1020
Db 961 TCAGACAACGCGCGCGCCCAAGATTTTCGACAAATTTTGGAAAATTTGAGCTGAAATGCG 1020
QY 1021 CAGAGTTTGGGATTAACGAAATGTTCGAGATTGACAGAACGACATCTGACATTTCTAGATT 1080
Db 1021 CAGAGTTTGGGATTAACGAAATGTTCGAGATTGACAGAACGACATCTGACATTTCTAGATT 1080
QY 1081 TGTGGAGGAAAATCTCTCTTTGTTTATAAGTCACTGTTTAAATTCGGAAGGAAATG 1140
Db 1081 TGTGGAGGAAAATCTCTCTTTGTTTATAAGTCACTGTTTAAATTCGGAAGGAAATG 1140
QY 1141 CAAAAAGTACGATACATTTGTTTCAAAATCTATGAGGAAAACAAAAATTTCAATGAGT 1200
Db 1141 CAAAAAGTACGATACATTTGTTTCAAAATCTATGAGGAAAACAAAAATTTCAATGAGT 1200
QY 1201 TTCCCCACCTACCACTAGTCAAAGTTTAAAGTGGAGCAAAAAGATAACGCTGTACCAATGG 1260
Db 1201 TTCCCCACCTACCACTAGTCAAAGTTTAAAGTGGAGCAAAAAGATAACGCTGTACCAATGG 1260
QY 1261 AACATCTTGAAGTTTATGAGAACCCACAAAGATACAAAGATCGAATTTGATCTGGTGATGC 1320
```

1261 AACATCTTGAAGTTCATGAGAGCCACAAAGATACAGAAATCGAATTCGATCTGGTATGC 1320
1321 AAGACAAGTTCTTAAAGCGAGCTACACGAAACCTCACGACTACAAAGAAATACCCCTAA 1380
1321 AAGACAAGTTCTTAAAGCGAGCTACACGAAACCTCACGACTACAAAGAAATACCCCTAA 1380
1381 AATGCTGAAAGAAATGGATTTCTTCTGAAAGAGCTAAATTTGTTGAAAGATTTGGAT 1440
1381 AATGCTGAAAGAAATGGATTTCTTCTGAAAGAGCTAAATTTGTTGAAAGATTTGGAT 1440
1441 TATGCTCCAAATCTCAGATGATCGAATGTCAGGAAAGGTTTGAAGAGCCAAATGCTTG 1500
1441 TATGCTCCAAATCTCAGATGATCGAATGTCAGGAAAGGTTTGAAGAGCCAAATGCTTG 1500
1501 TGAATAGTGTAAATGAACAAATTAATAATGACACCAAGTGAATTCGTGGATTTCAAGAAAC 1560
1501 TGAATAGTGTAAATGAACAAATTAATAATGACACCAAGTGAATTCGTGGATTTCAAGAAAC 1560
1561 AATTGAATGTGTTCCCGAAAGAACTTTGCTGCTGCTGTTTGTAGTCAACGAAACAG 1620
1561 AATTGAATGTGTTCCCGAAAGAACTTTGCTGCTGCTGTTTGTAGTCAACGAAACAG 1620
1621 CGGAAATCCATGCTTTAGAGAGAGACGAGTTGTTAAAGTTCTACACCGAACTAAATGGTG 1680
1621 CGGAAATCCATGCTTTAGAGAGAGACGAGTTGTTAAAGTTCTACACCGAACTAAATGGTG 1680
1681 GTTGCAAGTTCGTTGGAATACGAATTTGGTGCCAAATGAACAGAGAGCGCAATCTATTA 1740
1681 GTTGCAAGTTCGTTGGAATACGAATTTGGTGCCAAATGAACAGAGAGCGCAATCTATTA 1740
1741 TGTAACGCGGAGGAAATGAATATGCTTCTACAAATTTGTACACTAAATACCGGAA 1800
1741 TGTAACGCGGAGGAAATGAATATGCTTCTACAAATTTGTACACTAAATACCGGAA 1800
1801 TCGGTAGATTGAATAGCGGCAACGAAGCGAAGATATGTTGAAAGCTTTCCCGATA 1860
1801 TCGGTAGATTGAATAGCGGCAACGAAGCGAAGATATGTTGAAAGCTTTCCCGATA 1860
1861 AAGAACAAAGTCTTAATGTTTCAATTTATCAATTTCCAAACGACAACTGAATGCTTACGGTT 1920
1861 AAGAACAAAGTCTTAATGTTTCAATTTATCAATTTCCAAACGACAACTGAATGCTTACGGTT 1920
1921 TTGTGAAACATTTATTCGATCACACCATCGGTGTAGCTTAATCAGCATATTTCTTGAAA 1980
1921 TTGTGAAACATTTATTCGATCACACCATCGGTGTAGCTTAATCAGCATATTTCTTGAAA 1980
1981 CAGTCACAAAGCTTTGGCATCACTAAGGCAAGAGAAAGGATCAAAACGAAATTTCTATC 2040
1981 CAGTCACAAAGCTTTGGCATCACTAAGGCAAGAGAAAGGATCAAAACGAAATTTCTATC 2040
2041 AATTTGCAATTTGAAATCAACCGGAAATAGGAGGATTAACAGGAGCTTGACTGTCAG 2100
2041 AATTTGCAATTTGAAATCAACCGGAAATAGGAGGATTAACAGGAGCTTGACTGTCAG 2100
2101 AATTTGCAATTTGAAATCAACCGGAAATAGGAGGATTAACAGGAGCTTGACTGTCAG 2160
2101 AATTTGCAATTTGAAATCAACCGGAAATAGGAGGATTAACAGGAGCTTGACTGTCAG 2160
2161 ATGTTGGAATTTGATGTAACTCACTCAACCTCTACAGTGGAAATGATTTCTATAGCGG 2220
2161 ATGTTGGAATTTGATGTAACTCACTCAACCTCTACAGTGGAAATGATTTCTATAGCGG 2220
2221 CTGTAGTACGAGTATCAATCCAGGTGGAACTATCTATCGAAATATGATTTGACTCAAG 2280
2221 CTGTAGTACGAGTATCAATCCAGGTGGAACTATCTATCGAAATATGATTTGACTCAAG 2280
2281 AAGAAATGCTCCCGTGGAGCTGCTCATGCGGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAG 2340
2281 AAGAAATGCTCCCGTGGAGCTGCTCATGCGGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAG 2340
2341 AAGCAAAGTTCGTGAAATTTGCTCAGAGAAATTCGAGAAACCAACGAAATTCGAGCAACAG 2400

2341 AAGCAAAGTTCGTGAAATTTGCTCAGAGAAATTCGAGAAACCAACGAAATTCGAGCAACAG 2400
2401 CGCATATTTGTAGTCTATCGAGCGGAGTTAGCGATTCGAGAGATGCTACGTGTAGTCAATG 2460
2401 CGCATATTTGTAGTCTATCGAGCGGAGTTAGCGATTCGAGAGATGCTACGTGTAGTCAATG 2460
2461 ATGAGCTTCGATCTTTTAAAGAGCGAAGTAAACAAATTCATGTCGGAACGGGATGGAGAAG 2520
2461 ATGAGCTTCGATCTTTTAAAGAGCGAAGTAAACAAATTCATGTCGGAACGGGATGGAGAAG 2520
2521 ATCCAGAGCCGAGTACACGTTTCAATGATTCAGAAAAAGACAAATACACGATTCCTTTC 2580
2521 ATCCAGAGCCGAGTACACGTTTCAATGATTCAGAAAAAGACAAATACACGATTCCTTTC 2580
2581 GAAGAAATGGAAGAAAGATAGCCAGTGGTCAATAAAGATCTTACTCTGCTGAAACAGATG 2640
2581 GAAGAAATGGAAGAAAGATAGCCAGTGGTCAATAAAGATCTTACTCTGCTGAAACAGATG 2640
2641 TCGCTGTTGCTGCTGTTAAACAAATGGAGGAGGATATGAAGAAAGCAAGAAACCTGGAA 2700
2641 TCGCTGTTGCTGCTGTTAAACAAATGGAGGAGGATATGAAGAAAGCAAGAAACCTGGAA 2700
2701 TTGTGAACCCATCATCCGAAACAACTGGATTAATCTATCGTTTCGAAATACAAATTCG 2760
2701 TTGTGAACCCATCATCCGAAACAACTGGATTAATCTATCGTTTCGAAATACAAATTCG 2760
2761 ATTTTTCCTGTCATCTCATGTCGTTGCTGTTGATACATCTGTCGAGGACATACACATG 2820
2761 ATTTTTCCTGTCATCTCATGTCGTTGCTGTTGATACATCTGTCGAGGACATACACATG 2820
2821 TTATGTCATGACGATTAAGGAATGAGCCAGATGAAGTCTATGAATGACCTACGAGCTTG 2880
2821 TTATGTCATGACGATTAAGGAATGAGCCAGATGAAGTCTATGAATGACCTACGAGCTTG 2880
2881 CTTTTCCTCTGCTAGATGTGAAACCCCATCTCGTTGCTGTTCCGGTTCAATATGCTC 2940
2881 CTTTTCCTCTGCTAGATGTGAAACCCCATCTCGTTGCTGTTCCGGTTCAATATGCTC 2940
2941 ATTTTTCATGTCGAAAGGAAAGCTTTATCGAATCTACAGGAAACATACATCGGTG 3000
2941 ATTTTTCATGTCGAAAGGAAAGCTTTATCGAATCTACAGGAAACATACATCGGTG 3000
3001 ACTATGTCAGCAGCAGGACTCGACACGAAATCGAAATCTTCTCCAAATCTAAGCTGAAAT 3060
3001 ACTATGTCAGCAGCAGGACTCGACACGAAATCGAAATCTTCTCCAAATCTAAGCTGAAAT 3060
3061 ACCCTGGAATGTCGTTTCGCAATTAACATTTTGCAAAAGTGTGCGCCGTTTCAATCAAAATTTT 3120
3061 ACCCTGGAATGTCGTTTCGCAATTAACATTTTGCAAAAGTGTGCGCCGTTTCAATCAAAATTTT 3120
3121 TCAATTTGATATTTGATTTACTTTTAAAGCCCGGTTTCAAAATTCATTCATCCATG 3180
3121 TCAATTTGATATTTGATTTACTTTTAAAGCCCGGTTTCAAAATTCATTCATCCATG 3180
3181 ACTAAGCTTTTCAATTAATTTACTTTGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 3227
3181 ACTAAGCTTTTCAATTAATTTACTTTGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 3227

RESULT 4
US-10-645-746-1
; Sequence 1. Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UNY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20

;; PRIOR APPLICATION NUMBER: US 09/689,992
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/193,218
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: US 60/159,776
;; PRIOR FILING DATE: 1999-10-15
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 1
;; LENGTH: 3719
;; TYPE: DNA
;; ORGANISM: Caenorhabditis elegans
us-10-645-746-1

Query Match 80.6%; Score 2601; DB 9; Length 3719;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 3210; Conservative 0; Mismatches 0; Indels 509; Gaps 10;

QY 1 CAGCCACAAGTGTAGAAACATGCTCTCGAATTTTCCGGAATTCGGAAGGATTTTATC 60
Db 1 CAGCCACAAGTGTAGAAACATGCTCTCGAATTTTCCGGAATTCGGAAGGATTTTATC 60

QY 61 GTCATTCTCTCGATCC-----76
Db 61 GTCATTCTCTCGATCCGATATCAATTATTAGCAGCTATAGATATATAAGTTTGATA 120

QY 77 -----GGAGATGAATGCTTTCGAGGCCCACTGGTAATCGCAGCGCAAAATTC 125
Db 121 TTAATATTAAGGAGATGAATGCTTTCGAGGCCCACTGGTAATCGCAGCGCAAAATTC 180

QY 126 TAGGAGAAGAACTACTCTTTTGGTAAATGGTTCAGTTCTCCAGCAAAATTTAGCAT 185
Db 181 TAGGAGAAGAACTACTCTTTTGGTAAATGGTTCAGTTCTCCAGCAAAATTTAGCAT 240

QY 186 CGGGAATPACTACGAGTATGAAGTGAAATGACAAAGGAAGTATTGAATGAAACACGAGA 245
Db 241 CGGGAATPACTACGAGTATGAAGTGAAATGACAAAGGAAGTATTGAATGAAACACGAGA 300

QY 246 AAACCTTTCCAAAAAGACAGAAATTCCAAT-----277
Db 301 AAACCTTTCCAAAAAGACAGAAATTCCAATGTAAGTGTCTGTAAATTAGTCAAAACTA 360

QY 278 -----TCCGAGCTGCAAACTCTTCTGGCAACATCTTCGGCATGAGAAG 323
Db 361 ATTTTATTTTCACTCCGATCGTGCAGAACTCTTCTGGCAACATCTTCGGCATGAGAAG 420

QY 324 AAGCAGACAGATTTTATTTCTCGAAGACTATGTTTGTGATGAAAAGGACACTGTTTATAGT 383
Db 421 AAGCAGACAGATTTTATTTCTCGAAGACTATGTTTGTGATGAAAAGGACACTGTTTATAGT 480

QY 384 GTTTGTCGACTGAACACTGTGCAATCAAAAATGCTGTTTCGAGAAAGTAGTAAAAAAG 443
Db 481 GTTTGTCGACTGAACACTGTGCAATCAAAAATGCTGTTTCGAGAAAGTAGTAAAAAAG 540

QY 444 GATTCGGAGAAAAGATGAAAGGATTTGGAGAAAATCTTATACAAATGATACTT 503
Db 541 GATTCGGAGAAAAGATGAAAGGATTTGGAGAAAATCTTATACAAATGATACTT 600

QY 504 ACCTATCGTAAAAATTTTCACTTGAATTTAGTTCGAGAAAATCCGGAAGACGAGAA 563
Db 601 ACCTATCGTAAAAATTTTCACTTGAATTTAGTTCGAGAAAATCCGGAAGACGAGAA 660

QY 564 GCGAATCGGAGTTTACAAATTCCTGA-----588
Db 661 GCGAATCGGAGTTTACAAATTCCTGAAGGTTTATGAAAACACGCAATTTATAACAAACAAA 720

QY 589 -----AGATGTTATGACCCAGAAAGTTTCGCTACCGGCCCTTTTGTGAACGAGAGA 639
Db 721 TTAGCTTTTCAAGATGTTATGACCCAGAAAGTTTCGCTACCGGCCCTTTTGTGAACGAGAGA 780

QY 640 TTAAGT-----646
Db 781 TTAAGTGTGAGTTGCAATTAATAATAATCACTCACTCACTCACTCACTCACTCACTCACT 840

QY 647 -ACAAATTCGGAAAAATTTTGTGTACGATAATAATTCAAATTCCTCGAGTTCCTGAATCGT 705
Db 841 GACAAATTCGGAAAAATTTTGTGTACGATAATAATTCAAATTCCTCGAGTTCCTGAATCGT 900

QY 706 TTCAGGATCCAAACAGATTCGAAACATCTTAGAAGTAGACCAAGAAATCGAAGATCGATGGT 765
Db 901 TTCAGGATCCAAACAGATTCGAAACATCTTAGAAGTAGACCAAGAAATCGAAGATCGATGGT 960

QY 766 TTGGAAATTTACATTTGGAAATCAAGAAATTTGTCGATGTGAACTGTGCTCAATTTTGCA 825
Db 961 TTGGAAATTTACATTTGGAAATCAAGAAATTTGTCGATGTGAACTGTGCTCAATTTTGCA 1020

QY 826 -----TTGTGTAATAACT 838
Db 1021 GTAAGTTTGAGAAATTCGGATAAAAAATCATGTGATTTTGTGTAAGTTGTGTAATAACT 1080

QY 839 ATTCTACAATGCAACCGAAATGCTCTCTCTGGATATCTTCTCTAATTTGTGCAACCCCA 898
Db 1081 ATTCTACAATGCAACCGAAATGCTCTCTCTGGATATCTTCTCTAATTTGTGCAACCCCA 1140

QY 899 GTCGTGTAAACGATGTACGAAAGATCTTAAACAAACCTGATGCGGGAATAATGAC 958
Db 1141 GTCGTGTAAACGATGTACGAAAGATCTTAAACAAACCTGATGCGGGAATAATGAC 1200

QY 959 AATCAGACAAAGCCGCGCCCAAGAAATTCGACAAATTTAGGAAATTTGAGCTGAAATG 1018
Db 1201 AATCAGACAAAGCCGCGCCCAAGAAATTCGACAAATTTAGGAAATTTGAGCTGAAATG 1260

QY 1019 CGCAGAAATTTGGGATACCAAT-----1042
Db 1261 CGCAGAAATTTGGGATACCAATTTAGTAAATTTATTCAAACAATTAATATACAAAT 1320

QY 1043 -----GTCCAGATTCACAGACGACATCTGACATTTCTAGATTTGTGCGCAGGAAA 1092
Db 1321 TGATTTTCAGTTCGAGATTCGAGAAACGACATCTGACATTTCTAGATTTGTGCGCAGGAAA 1380

QY 1093 ACTCTCTGTTTATAAGTCACTGGTAAATCGGACAGAGGAGAAATTCGAAAAAAGTACG 1152
Db 1381 ACTCTCTGTTTATAAGTCACTGGTAAATCGGACAGAGGAGAAATTCGAAAAAAGTACG 1440

QY 1153 ATACTACATTTTCAAAATCTATGAGGAAACAAAAAGTTTCATTTGAGTTTCCCACTAC 1212
Db 1441 ATACTACATTTTCAAAATCTATGAGGAAACAAAAAGTTTCATTTGAGTTTCCCACTAC 1500

QY 1213 CACTAGTCAAGTTTAAAGTGAGGACAAAGAAATACGCTGTACCAATCGAAATCTCTCAAG 1272
Db 1501 CACTAGTCAAGTTTAAAGTGAGGACAAAGAAATACGCTGTACCAATCGAAATCTCTCAAG 1560

QY 1273 TTCAATGAGAGCCCAAAAGATACAAAGATTCGAAATTCGATCTGTTGATGCAAGACAAAGTTTC 1332
Db 1561 TTCAATGAGAGCCCAAAAGATACAAAGATTCGAAATTCGATCTGTTGATGCAAGACAAAGTTTC 1620

QY 1333 TAAAGCGAGTACAGAAACCTCAGCATACAAAGAAATACCTTAAATAATGCTGAAAG 1392
Db 1621 TAAAGCGAGTACAGAAACCTCAGCATACAAAGAAATACCTTAAATAATGCTGAAAG 1680

QY 1393 AATTGATTTCTCTCTGAGAGCTAAAATTTTGTGCAAGATTTGGATTTATGCTCCAAAC 1452
Db 1681 AATTGATTTCTCTCTGAGAGCTAAAATTTTGTGCAAGATTTGGATTTATGCTCCAAAC 1740

QY 1453 TTCAGATGATCGAAATGTCAGGAAAGGTTTGAAGAGCCAAATGCTTTGTAATAGTGTAA 1512
Db 1741 TTCAGATGATCGAAATGTCAGGAAAGGTTTGAAGAGCCAAATGCTTTGTAATAGTGTAA 1800

QY 1513 ATGAAACAAATTAATAATGACACGATGATTCGTTGATTTTCAAGAAAAACAATTAATGTTGG 1572
Db 1801 ATGAAACAAATTAATAATGACACGATGATTCGTTGATTTTCAAGAAAAACAATTAATGTTGG 1860

QY 1573 TTTCCGAAAAAGAACTTTCTGCTGCTGTTTCTGCTAGTCAACGAAACAGCGGGAATCCAT 1632
Db 1861 TTTCCGAAAAAGAACTTTCTGCTGCTGTTTCTGCTAGTCAACGAAACAGCGGGAATCCAT 1920

QY 1633 GCTTAGAAGAGACGAGTTGT----- 1654
 Db 1921 GCTTAGAGAGAGACGAGTTGTAAAGTGTCTTACGTAGATATTCCGAAATATTTTC 1980
 QY 1655 --TAAGTTCTACACCGAACTAAATGGTGGTGTGCAAGTTCGGTGAATACGAAATGGTGCC 1712
 Db 1981 AGTAAGTTCTACACCGAACTAAATGGTGGTGTGCAAGTTCGGTGAATACGAAATGGTGCC 2040
 QY 1713 AATGAANAACAGAGGAGCGCAATCTATATGTACGACGCGACGAAATGAATATGCC---- 1769
 Db 2041 AATGAANAACAGAGGAGCGCAATCTATATGTACGACGCGACGAAATGAATATGCCGTA 2100
 QY 1770 -----TTCTACAAAAATTTGACAC 1788
 Db 2101 AGTTTCAGAAAAATTTGAAAGTTTTTAAATATCATATTACAGTTCTACAAAAATTTGACAC 2160
 QY 1789 TAAATACCGGAATCGGTAGATTGAAATAGCCGCAACAGAAGCGAAGAAATATGTTTGAAC 1848
 Db 2161 TAAATACCGGAATCGGTAGATTGAAATAGCCGCAACAGAAGCAATATGTTTGAAC 2220
 QY 1849 GTCTTCCGATAAAGAACAAAAAGTCTTAATGTTCAATTATCATTTTCCAAAACGACACTGA 1908
 Db 2221 GTCTTCCGATAAAGAACAAAAAGTCTTAATGTTCAATTATCATTTTCCAAAACGACACTGA 2280
 QY 1909 ATGCTTACCGTTTGTGAAACATTAATGCGATCACACCATCGGTGTAGCTAAATCAGCATA 1968
 Db 2281 ATGCTTACCGTTTGTGAAACATTAATGCGATCACACCATCGGTGTAGCTAAATCAGCATA 2340
 QY 1969 TTACTTCTGAAACAGTCAAAAAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2028
 Db 2341 TTACTTCTGAAACAGTCAAAAAAGCTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2400
 QY 2029 GAATTTTCTATCAAAATTCGAATTCGAAATCAACGCGAAATTAGAGGTATTAAACGAGGC 2088
 Db 2401 GAATTTTCTATCAAAATTCGAATTCGAAATCAACGCGAAATTAGAGGTATTAAACGAGGC 2460
 QY 2089 TTGACTGTGTCAGAAATTCGAGAAATATCACAGAGAAAGAAAGACGCGAAACAAATGC 2148
 Db 2461 TTGACTGTGTCAGAAATTCGAGAAATATCACAGAGAAAGAAAGACGCGAAACAAATGC 2520
 QY 2149 CATTAACATATGATGTGGAATGTGAATCACTCAACCTCTCACTAGTGGAAATGTATT 2208
 Db 2521 CATTAACATATGATGTGGAATGTGAATCACTCAACCTCTCACTAGTGGAAATGTATT 2580
 QY 2209 ATCTATACCGCTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCGAAATATGA 2268
 Db 2581 ATCTATACCGCTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCGAAATATGA 2640
 QY 2269 TTGTGACTCAAGAGAAATGTCCTCCGGTGAGCGTGCAGTGGCTCATGGACGCGAAAGAA 2328
 Db 2641 TTGTGACTCAAGAGAAATGTCCTCCGGTGAGCGTGCAGTGGCTCATGGACGCGAAAGAA 2700
 QY 2329 CAGATATTTTGGAGCAAAAGTTGCGTGAATTCGTCAGAGAAATTCGCAGAA----- 2378
 Db 2701 CAGATATTTTGGAGCAAAAGTTGCGTGAATTCGTCAGAGAAATTCGCAGAGAAATTCGTCAG 2760
 QY 2379 -----AACACGAC 2387
 Db 2761 TTGAGTATTAAAGATCTCGGATTTTTTAATTTTTTTTGTAAACTTTTCAGAACACGAC 2820
 QY 2388 AATCGAGCACCGCGCATATTGTAGTCTATCGAGCGGAGTTAGCGATTCCGAGATGCTA 2447
 Db 2821 AATCGAGCACCGCGCATATTGTAGTCTATCGAGCGGAGTTAGCGATTCCGAGATGCTA 2880
 QY 2448 CCGTTTGTATGATGAGCTTCGATCTTTTAAAGCGAAGTAAAAACAATTCATGTCGAA 2507
 Db 2881 CCGTTTGTATGATGAGCTTCGATCTTTTAAAGCGAAGTAAAAACAATTCATGTCGAA 2940
 QY 2508 CCGGATGGAGAGATCCAGCGCGAGTACAGTTTCATTTGATTCAGAAAGACACAT 2567
 Db 2941 CCGGATGGAGAGATCCAGCGCGAAGTACAGTTTCATTTGATTCAGAAAGACACAT 3000
 QY 2568 ACACGATTCTTCGAAGAATGGAAAAAGATAAGCCAGTGGTCAATAAAGATCTTACTCCT 2627

Db 3001 ACACCATGCTTCGAAGAATGGAAAAAGATAAGCCAGTGGTCAATAAAGATCTTACTCCT 3060
 QY 2628 GCTGAAAACAGATGTCGCTGTTGCTGCTGTGTAAACAATGGGAGGAGATATGAAAGAAAGC 2687
 Db 3061 GCTGAAAACAGATGTCGCTGTTGCTGCTGTGTAAACAATGGGAGGAGATATGAAAGAAAGC 3120
 QY 2688 AAGGAAACTGGAAATTTGTGAACCCATCCGGAACAACTGTGGAATAAACTTATCGTTTCG 2747
 Db 3121 AAGGAAACTGGAAATTTGTGAACCCATCATCCGGAACAACTGTGGAATAAACTTATCGTTTCG 3180
 QY 2748 AAATACAAATTCGAATTTTTTTTGTGGCATCTCATCATGTGTGCTTGGTATCATCTCGTCCA 2807
 Db 3181 AAATACAAATTCGAATTTTTTTTGTGGCATCTCATCATGTGTGCTTGGTATCATCTCGTCCA 3240
 QY 2808 GGACATTACACTGTTATGTATGACGATAAAGAAATGAGCAAGATGAAGTCTAT 2861
 Db 3241 GGACATTACACTGTTATGTATGACGATAAAGAAATGAGCAAGATGAAGTCTATGTAAGC 3300
 QY 2862 ----- 2861
 Db 3301 GTTTTGAATAGCAGTTAGCGATTTTAGGATTTTGTAAATCCGCATATATGTTATTATAAAA 3360
 QY 2862 -----AAAATGACCTACGGACTTGTCTTCTCTGCTAGATGTGAAAAACCCAT 2911
 Db 3361 AATGTTTCAGAAAAATGACCTACGGACTTGTCTTCTCTGCTAGATGTGAAAAACCCAT 3420
 QY 2912 CTCGTTGCTGTTCCGGTTCTATATGCTCATTTATCATGTGAAAAAGCGAAAGAGCTTTA 2971
 Db 3421 CTCGTTGCTGTTCCGGTTCTATATGCTCATTTATCATGTGAAAAAGCGAAAGAGCTTTA 3480
 QY 2972 TCGAACTTACAGGAACATTAATCATCGGTGACTATGCACAGCCAGCAGCTCGACACGAAT 3031
 Db 3481 TCGAACTTACAGGAACATTAATCATCGGTGACTATGCACAGCCAGCAGCTCGACACGAAT 3540
 QY 3032 GGAAACATTTCTCCAAACTTAACGTAAGTACCTCGGAATGTCGTTGCGCATAAACATTTGC 3091
 Db 3541 GGAAACATTTCTCCAAACTTAACGTAAGTACCTCGGAATGTCGTTGCGCATAAACATTTGC 3600
 QY 3092 AAAAGTGTGCGCGTTTCAATCAAAATTTTCAATTTGATAGATATTGTACTTTTTTTT 3151
 Db 3601 AAAAGTGTGCGCGTTTCAATCAAAATTTTCAATTTGATAGATATTGTACTTTTTTTT 3660
 QY 3152 AAAAGTGTGCGCGTTTCAAAATTTCAATTCATGACTAAACGTTTTCATAAAATTTGAAATTT 3210
 Db 3661 AAAAGTGTGCGCGTTTCAAAATTTCAATTCATGACTAAACGTTTTCATAAAATTTGAAATTT 3719

RESULT 5

US-10-645-735-1
 ; Sequence 1, Application US/10645735
 ; Publication No. US20050100913A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mello, Craig C.
 ; APPLICANT: Tabara, Hiroaki
 ; APPLICANT: Grishok, Alla
 ; APPLICANT: Fire, Andrew
 ; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
 ; FILE REFERENCE: INTERFERENCE
 ; CURRENT APPLICATION NUMBER: US/10/645,735
 ; PRIOR FILING DATE: 2003-08-20
 ; PRIOR APPLICATION NUMBER: US/09/689,992A
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: US 60/193,218
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: US 60/159,776
 ; PRIOR FILING DATE: 1999-10-15
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3719
 ; TYPE: DNA

i: ORGANISM: Caenorhabditis elegans
US-10-645-735-1

Query Match 80.6%; Score 2601; DB 10; Length 3719;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 3210; Conservative 0; Mismatches 0; Indels 509; Gaps 10;

```
QY 1 CAGCCAGAGTGTGTAAGCAATGCTCCCTCGAATTTTCCCGAATTCGGAAGGATTTTATC 60
DB 1 CAGCCAGAGTGTGTAAGCAATGCTCCCGAATTTTCCCGAATTCGGAAGGATTTTATC 60
QY 61 GTCATTCTCTCGATCC-----76
DB 61 GTCATTCTCTCGATCCGATGATGATCAATATTAGCAGCTATAAGATATATAAGTTTGATA 120
QY 77 -----GGAGATGAAATGCTTTCGAGGCCCACTGGTAAATCGCAGCGCAAAATTC 125
DB 121 TTAATAATTATAGAGATGAAATGGCTTTCGAGGCCCACTGGTAAATCGCAGCGCAAAATTC 180
QY 126 TATGAGAGAGAGTACTTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTTACGAT 185
DB 181 TATGAGAGAGAGTACTTCTTTTGGTAAATGGTTCAAGTTCTCCAGCAAAATTTACGAT 240
QY 186 CGGGAATACCTACGAGTATGAAGTAAATGCAAAAGGAAGTATTGAAATAGAAACACGAGA 245
DB 241 CGGGAATACCTACGAGTATGAAGTAAATGCAAAAGGAAGTATTGAAATAGAAACACGAGA 300
QY 246 AAACCTTTCCAAAAGACAGAAATTCGAAT-----277
DB 301 AAACCTTTCCAAAAGACAGAAATTCGAATGTAAGTGTCTGTAAATTAGTCAAAACTA 360
QY 278 -----TCCGATCGTGCAAACTCTTCTGGCAACATCTTCGGCATGAGAAG 323
DB 361 ATTTTATTTTTCAGTCCCGATCGTGCAAACTCTTCTGGCAACATCTTTCGGCATGAGAAG 420
QY 324 AAGCAGACAGATTTTATCTCGAAGACTATGTTTTTGTGATGAAAGGACACTGTTTATAGT 383
DB 421 AAGCAGACAGATTTTATCTCGAAGACTATGTTTTTGTGATGAAAGGACACTGTTTATAGT 480
QY 384 GTTTGTGCACTGAACACTGTGCATCAAAATGCTGGTTTCGGAGAAAGTAGTAAAAAAG 443
DB 481 GTTTGTGCACTGAACACTGTGCATCAAAATGCTGGTTTCGGAGAAAGTAGTAAAAAAG 540
QY 444 GATTCGGAGAAAAGATGAAAAGGATTTGGAGAAAAAATCTTTATACAAATGATACTT 503
DB 541 GATTCGGAGAAAAGATGAAAAGGATTTGGAGAAAAAATCTTTATACAAATGATACTT 600
QY 504 ACCTATCGTAAAAATTTTCACTGAACTTTTGTAGTCGAGAAAATCCGGAAGAACGAGAA 563
DB 601 ACCTATCGTAAAAATTTTCACTGAACTTTTGTAGTCGAGAAAATCCGGAAGAACGAGAA 660
QY 564 GCGAATCGGAGTTACAAATTCCTGA-----588
DB 661 GCGAATCGGAGTTACAAATTCCTGAAGGTTTATGAAAAACACGCAATTATACAAACAAA 720
QY 589 -----AGAAATGTTATGACCCAGAAAGTTTCGCTACGCGCCTTTTGTGAAACGAGAGA 639
DB 721 TTAGCTTTTCAGAAATGTTATGACCCAGAAAGTTTCGCTACGCGCCTTTTGTGAAACGAGAGA 780
QY 640 TTAAGT-----646
DB 781 TTAAGTGTGAGTTGCAATTAATTAATAATATCACTCACTCAATTAATTAATTTTAA 840
QY 647 -ACAAATTCGCGAAAAATTTTGTAGCAATAAATTAATTTCTGCGAGTTTCTTGAATCGT 705
DB 841 GACAAATTCGCGAAAAATTTTGTAGCAATAAATTAATTTCTGCGAGTTTCTTGAATCGT 900
QY 706 TTCACGATCAAAACAGATTCGAAACAATCATATTAGAAGTAGCAACAAGATCGAAGCATGGT 765
DB 901 TTCACGATCAAAACAGATTCGAAACAATCATATTAGAAGTAGCAACAAGATCGAAGCATGGT 960
QY 766 TTGGAATTTACATTGGAATCAAGAAATTTGTCGATGGTGAACCTGCTCAATTTTGCBA 825
DB 825
```

```
DB 961 TTGGAATTTACATTGGAATCAAGAAATTTGTTGATGTTGAACTGTGCTCAATTTTGCA 1020
QY 826 -----TTGTTCGATAAACT 838
DB 1021 GTAAGTTTGAGAAACTCGGATAAAAAATCATGTGATTTTGTGGAAGTTGTTCGATAAACT 1080
QY 839 ATTCTACAATGCAACCGAAATGTCTCTTCTGGATTTATCTTCTCTTAATTTGTTCGACCCCA 898
DB 1081 ATTCTACAATGCAACCGAAATGTCTCTTCTGGATTTATCTTCTCTTAATTTGTTCGACCCCA 1140
QY 899 GTCGTGTAAACGATGATGTACGAAAGATCTTAAACAAAACTGATGCGGGAATAATGAC 958
DB 1141 GTCGTGTAAACGATGATGTACGAAAGATCTTAAACAAAACTGATGCGGGAATAATGAC 1200
QY 959 AATCAGACAAAGCGCGGCCCAAGAAATTCGACAAATTTATGGAATAATTTGAAGCTGAAATG 1018
DB 1201 AATCAGACAAAGCGCGGCCCAAGAAATTCGACAAATTTATGGAATAATTTGAAGCTGAAATG 1260
QY 1019 CGCAGAAAGTTTGGGATAACGAAT-----1042
DB 1261 CGCAGAAAGTTTGGGATAACGAATTTAGTTTAAATTTATTCAAACAATTTAATATACAAAT 1320
QY 1043 -----GTCGAGATTGACAGAACGACATCTGACATTTCTAGATTTGTGCGAGGAAA 1092
DB 1321 TGATTTTTCAGGTGAGATTTGACAGAACGACATCTGACATTTCTAGATTTGTGCGAGGAAA 1380
QY 1093 ACTCTCTGTTTATAAAGTCACTGGTAAATCGGACAGAGAGAAATGCAAAAAGATGACG 1152
DB 1381 ACTCTCTGTTTATAAAGTCACTGGTAAATCGGACAGAGAGAAATGCAAAAAGATGACG 1440
QY 1153 ATACTACATTTTCAAAATCTATGAGGAAAAACAAAAGTTTCATTTGAGTTTCCCACTAC 1212
DB 1441 ATACTACATTTTCAAAATCTATGAGGAAAAACAAAAGTTTCATTTGAGTTTCCCACTAC 1500
QY 1213 CACTAGTCAAAAGTTTAAAGTGGAGCAAAAGAAATACCTGTACCAATGGAACATCTTCAAG 1272
DB 1501 CACTAGTCAAAAGTTTAAAGTGGAGCAAAAGAAATACCTGTACCAATGGAACATCTTCAAG 1560
QY 1273 TTCATGAGAGGCCAACAAAGATACAAGAAATCGAAATTCGATCTGTGATGCAAGACAAAGTTTC 1332
DB 1561 TTCATGAGAGGCCAACAAAGATACAAGAAATCGAAATTCGATCTGTGATGCAAGACAAAGTTTC 1620
QY 1333 TAAAGCGAGGTACACGAAACCTCACGACTACAAGAAATAACCTTAATAAATGCTGAAG 1392
DB 1621 TAAAGCGAGGTACACGAAACCTCACGACTACAAGAAATAACCTTAATAAATGCTGAAG 1680
QY 1393 AATTGGAATTTCTCTTCTGAAGAGCTAAAATTTTCTGCAAGATTTTGAATATGCTCCAAAC 1452
DB 1681 AATTGGAATTTCTCTTCTGAAGAGCTAAAATTTTCTGCAAGATTTTGAATATGCTCCAAAC 1740
QY 1453 TTCAGATGATCGAATGTCAGGAAAGGTTTGAAGAGGCAATGCTTGTGAATAGTGTAA 1512
DB 1741 TTCAGATGATCGAATGTCAGGAAAGGTTTGAAGAGGCAATGCTTGTGAATAGTGTAA 1800
QY 1513 ATGAACAAATTTAAATGACACGAGTATTCGTTGATTCAGAGAAACAAATTTGAATGTGG 1572
DB 1801 ATGAACAAATTTAAATGACACGAGTATTCGTTGATTCAGAGAAACAAATTTGAATGTGG 1860
QY 1573 TTCCCGAAAAAGAACTTTCTGCTGCTGTTTTTGTAGTCAACGAAACAGCGGGAATCCAT 1632
DB 1861 TTCCCGAAAAAGAACTTTCTGCTGCTGTTTTTGTAGTCAACGAAACAGCGGGAATCCAT 1920
QY 1633 GCTTAGAAGAGAACGACGTTGT-----1654
DB 1921 GCTTAGAAGAGAACGACGTTGTGTAGTGTTCACGTAGATTTATTCGGAATATTTTC 1980
QY 1655 --TAAGTCTTACACCGAACTAAATTTGTTGTTGTCGCAAGTTTCGTTGGAATATACGAATGTTGTC 1712
DB 1981 AGTAAGTTCTTACACCGAACTAAATTTGTTGTTGTCGCAAGTTTCGTTGGAATATACGAATGTTGTC 2040
QY 1713 AATGAAAAACAGAGGAGCGCAATCTATTATGTACGACGCGACGAAATGAATATGTC--- 1769
DB 2041 AATGAAAAACAGAGGAGCGCAATCTATTATGTACGACGCGACGAAATGAATATGTCGTA 2100
```

QY 1770 -----TTCTCAAAAAATTGTACAC 1788
 DB AGTTTCAGAAAAATTGAAAGTTTTTAAATATCATATTTACAGTTCTACAAAAAATTGTACAC 2160
 QY TAAATACCGGAATCGGTAGATTGAAATAGCGCAACAGAGCGGAAGAAATATGTTTGAAC 1848
 DB TAAATACCGGAATCGGTAGATTGAAATAGCGCAACAGAGCGGAAGAAATATGTTTGAAC 2220
 QY GTCTTCCCGATTAAGAAACAAAAAGTCTTAATGTTCATTATCATTTTCCAAACGCAACTGA 1908
 DB GTCTTCCCGATTAAGAAACAAAAAGTCTTAATGTTCATTATCATTTTCCAAACGCAACTGA 2280
 QY ATGCTTACCGTTTTGTGAAACATTATTTGCGATCACACCATCGGTGTAGCTAATCAGCAT 1968
 DB ATGCTTACCGTTTTGTGAAACATTATTTGCGATCACACCATCGGTGTAGCTAATCAGCAT 2340
 QY TTACTTCTGAAACAGTCAACAAAGCTTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2028
 DB TTACTTCTGAAACAGTCAACAAAGCTTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2400
 QY GAATTTTCTATCAAAATTGCATTGAAATCAACCGGAAATTTAGAGGTATTTAACCGAGGC 2088
 DB GAATTTTCTATCAAAATTGCATTGAAATCAACCGGAAATTTAGAGGTATTTAACCGAGGC 2460
 QY TTGACTGGTCAGAAATTGAGAAATATCACAGAGAAAGAAAGAGACGCGAAACCATGC 2148
 DB TTGACTGGTCAGAAATTGAGAAATATCACAGAGAAAGAAAGAGACGCGAAACCATGC 2520
 QY CATTAACATATGATGTGAAATTGATGTAACCTCATCCAACTCTTACAGTGAATTTGATT 2208
 DB CATTAACATATGATGTGAAATTGATGTAACCTCATCCAACTCTTACAGTGAATTTGATT 2580
 QY ATTCTATACCGCTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCTATCGAAATATGA 2268
 DB ATTCTATACCGCTGTAGTAGCGAGTATCAATCCAGGTGGAATCTATCTATCGAAATATGA 2640
 QY TTGTGACTCAAGAGATGTCGTCGGTGAGGTGCAGTGGCTCATGACCGGAAGAA 2328
 DB TTGTGACTCAAGAGATGTCGTCGGTGAGGTGCAGTGGCTCATGACCGGAAGAA 2700
 QY CAGATATTTTGAAGCAAAAGTTGCGTGAATTTGCTCAGAGAAATTCGAGAA----- 2378
 DB CAGATATTTTGAAGCAAAAGTTGCGTGAATTTGCTCAGAGAAATTCGAGAGTGTGTC 2760
 QY 2379 -----AACACGAC 2387
 DB TTGAGTATTTAAAGATCTCTGGATTTTAAATTTTTTTTGTAAACTTTTCAGAAACACGAC 2820
 QY AATCGAGCACCGCATATTTAGTCTATCGAGCGGAGTTAGCGATTTCGGAGATGCTA 2447
 DB AATCGAGCACCGCATATTTAGTCTATCGAGCGGAGTTAGCGATTTCGGAGATGCTA 2880
 QY CGTGTAGTCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAAATTCATGTCGAA 2507
 DB CGTGTAGTCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAAATTCATGTCGAA 2940
 QY CGGGATGGAAGAGATCCAGAGCGGAAGTACAGTTTCATTGTGATTCAGAAAGACACAAT 2567
 DB CGGGATGGAAGAGATCCAGAGCGGAAGTACAGTTTCATTGTGATTCAGAAAGACACAAT 3000
 QY ACACGATTCCTCGAAGATGGAAGAAAGATAGCCAGTGGTCAATAAAGATCTTACTCCT 2627
 DB ACACGATTCCTCGAAGATGGAAGAAAGATAGCCAGTGGTCAATAAAGATCTTACTCCT 3060
 QY GCTGAAACAGATGTCCTGCTGTGCTGTTTAAACATGGGAGGAGGATATGAAGAAAGC 2687
 DB GCTGAAACAGATGTCCTGCTGTGCTGTTTAAACATGGGAGGAGGATATGAAGAAAGC 3120
 QY AAGAAACCTGGAATTTGTGAACCCATCATCCGGAACAACTGTGTGATTAACCTTATCGTTTCG 2747
 DB AAGAAACCTGGAATTTGTGAACCCATCATCCGGAACAACTGTGTGATTAACCTTATCGTTTCG 3180

QY 2748 AAATACAAATTCGATTTTCTTGGCATCTCATGTTGTCCTTGGTATACATCTCGTCCA 2807
 DB AAATACAAATTCGATTTTCTTGGCATCTCATGTTGTCCTTGGTATACATCTCGTCCA 3240
 QY GGACATTTACACTGTTTATGTATGACGATTAAGGAATGAGCAAGATGAAGTCTAT 2861
 DB GGACATTTACACTGTTTATGTATGACGATTAAGGAATGAGCAAGATGAAGTCTATGTAAGC 3300
 QY 2862 ----- 2861
 DB GTTTTGAATFAGCAGTTAGCGATTTTAGGATTTTGTAAATCCGCATATAGTTATTAATAAAA 3360
 QY 2862 -----AAATGACCTACGACTTGTCTTCTCTGTAGATGTCGAAACCCAT 2911
 DB AATGTTTCAGAAATGACCTACGACTTGTCTTCTCTGTAGATGTCGAAACCCAT 3420
 QY CTCGTGTGCTGTTCCGGTTCATTATGCTCATTTATCATGTGAAAAAGCGAAGAGCTTTA 2971
 DB CTCGTGTGCTGTTCCGGTTCATTATGCTCATTTATCATGTGAAAAAGCGAAGAGCTTTA 3480
 QY TCGAATTTTACAAGGAACATTATCATCGTGACTATGCACAGCCACGACTCGACACGAAT 3031
 DB TCGAATTTTACAAGGAACATTATCATCGTGACTATGCACAGCCACGACTCGACACGAAT 3540
 QY GGAACATTTTCTCCAACTAAACGTGAAGTACCCTCGAATGTCGTTCCGATTAACATTTTGC 3091
 DB GGAACATTTTCTCCAACTAAACGTGAAGTACCCTCGAATGTCGTTCCGATTAACATTTTGC 3600
 QY AAAAGTGTGCGCGTTTCAATCAAAATTTTCAATTTTGTAGATTTGTACTTTTTTTT 3151
 DB AAAAGTGTGCGCGTTTCAATCAAAATTTTCAATTTTGTAGATTTGTACTTTTTTTT 3660
 QY AAAGCCCGTTTCAAAATTTTCAATGACTAAAGTTTTCATTAATTTTCAATTTTGAATTT 3210
 DB AAAGCCCGTTTCAAAATTTTCAATGACTAAAGTTTTCATTAATTTTCAATTTTGAATTT 3719

RESULT 6

US-11-144-985-1
 ; Sequence 1, Application US/11144985
 ; Publication No. US20060024798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mello, Craig C.
 ; APPLICANT: Tabara, Hiroaki
 ; APPLICANT: Grishok, Alla
 ; APPLICANT: Fire, Andrew
 ; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
 ; FILE REFERENCE: UMY-052CN
 ; CURRENT APPLICATION NUMBER: US/11/144,985
 ; PRIOR FILING DATE: 2005-06-03
 ; PRIOR APPLICATION NUMBER: US 09/689,992
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: US 60/193,218
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: US 60/159,776
 ; PRIOR FILING DATE: 1999-10-15
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3719
 ; TYPE: DNA
 ; ORGANISM: Caenorhabditis elegans
 US-11-144-985-1

Query Match 80.6%; Score 2601; DB 16; Length 3719;
 Best Local Similarity 86.3%; Pred. No. 0;
 Matches 3210; Conservative 0; Mismatches 0; Indels 509; Gaps 10;
 QY 1 CAGCCACAAAGTGATGAAACATGTCCTCGAATTTTCCGAAATTTGGAAGAGGATTTATC 60
 DB 1 CAGCCACAAAGTGATGAAACATGTCCTCGAATTTTCCGAAATTTTCCGAAATTTGGAAGAGGATTTATC 60

Qy	61	GTCA	TTCTCTCGATCC	-----	76	
Db	61	GTCA	TTCTCTCGATCCGCTATGATCAA	TTATTAGCAGCTATAAGATAATAAGTTTGATA	120	
Qy	77	-----	GGAGATCAAAATGGCTTCGGAGGCCACTG	TAAATGCGACGGCAAAATTC	125	
Db	121	TTAA	TATATATAGGAGATGA	AAATGGCTTCGGAGGCCACTG	TAAATGCGACGGCAAAATTC	180
Qy	126	TATG	AGAAAGTAGTACTTCTTTTGGTAAATTTGGT	TCAAGTTCTCCAGCAAAATTTACGAT	185	
Db	181	TATG	AGAAAGTAGTACTTCTTTTGGTAAATTTGGT	TCAAGTTCTCCAGCAAAATTTACGAT	240	
Qy	186	CGG	AACTACTACGAGTAGTGAAGTGA	AAATGACAAAGGAAGTATTGGAATAGAAAAACGAGA	245	
Db	241	CGG	AACTACTACGAGTAGTGAAGTGA	AAATGACAAAGGAAGTATTGGAATAGAAAAACGAGA	300	
Qy	246	AAAC	CTTTCCCAAAAAGACAGAAATTC	CAAT-----	277	
Db	301	AAAC	CTTTCCCAAAAAGACAGAAATTC	CAATGTAAGTGTGTAATTTAGTCAAAACTA	363	
Qy	278	-----	TCCCGATCGTGCAAAAATCTCTTCTGGCAACATCT	TCGGCATGAGGAAG	323	
Db	361	ATTT	TATTTTTCAGTCCCGATCGTGCAAAAATCTCTTCTGGCAACATCT	TCGGCATGAGGAAG	420	
Qy	324	AAG	CAGACAGATTTTATTTCTCGAAGACTATGTTTTTGATG	AAAAAGGACACTGTTTATAGT	383	
Db	421	AAG	CAGACAGATTTTATTTCTCGAAGACTATGTTTTTGATG	AAAAAGGACACTGTTTATAGT	480	
Qy	384	GTTT	GTGCACTGAACACTGTGCACATCA	AAATGCTGTTTTCGGAGAAAGTAGTAAAAAAG	443	
Db	481	GTTT	GTGCACTGAACACTGTGCACATCA	AAATGCTGTTTTCGGAGAAAGTAGTAAAAAAG	540	
Qy	444	GATT	CGGAGAAAAAGATGAAAGGATTTGGAGAAAAA	AAATCTTTATACACATGATCTT	503	
Db	541	GATT	CGGAGAAAAAGATGAAAGGATTTGGAGAAAAA	AAATCTTTATACACATGATCTT	600	
Qy	504	ACCT	ATCGTAAAAAATTTCACTCGAACCTTTAGT	CGAGAAAAATCGGAAAAAGACAGAA	563	
Db	601	ACCT	ATCGTAAAAAATTTCACTCGAACCTTTAGT	CGAGAAAAATCGGAAAAAGACAGAA	660	
Qy	564	GCG	AATCGGAGTTACAAATTCCTGA	-----	588	
Db	661	GCG	AATCGGAGTTACAAATTCCTGAAGGTTTTATG	AAAAAACACGCAATTTATRACAAAACAAA	720	
Qy	589	-----	AGAAATGTTATGACCCAGAAAGTTTCGCTAG	CGGCCCTTTTGTGAACACGAGGA	639	
Db	721	TTAG	CTTTCAGAAATGTTATGACCCAGAAAGTTTCGCTAG	CGGCCCTTTTGTGAACACGAGGA	780	
Qy	640	TTAA	AGT-----	-----	646	
Db	781	TTAA	AGTGTGAGTTGCAATATAATAATAATCACTCA	CACTCATTTATATATTTTAA	840	
Qy	647	-ACA	TTTCGGAAAAATTTTGTGTACGATAATAATTC	AAATTCGCTCGGAGTTCTCGAATCGT	705	
Db	841	GACA	AAATTCGGAAAAATTTTGTGTACGATAATAATAATTC	AAATTCGCTCGGAGTTCTCGAATCGT	900	
Qy	706	TTCA	CGATCCAAAACAGATTCGAAACAAATCATTTAG	AAGTAGCACCAAGAAATCGAAGCATGGT	765	
Db	901	TTCA	CGATCCAAAACAGATTCGAAACAAATCATTTAG	AAGTAGCACCAAGAAATCGAAGCATGGT	960	
Qy	766	TTG	AAATTTTACATTTGGAATCAAGAAATTTGTT	TCGATGGTGAACCTGTGCTCAATTTTGC	825	
Db	961	TTG	AAATTTTACATTTGGAATCAAGAAATTTGTT	TCGATGGTGAACCTGTGCTCAATTTTGC	1020	
Qy	826	-----	-----	-----	TTGTGATAAACT	838
Db	1021	GTA	AGTTTGAAAACTCGATAAAAAAATCATGT	GATTTTGTGGAAGTTGTGATAAACT	1080	
Qy	839	ATT	CTCAATGCAACCGAAAAATGCTCTCTTCG	GAATATCTTCTCCTTAATTCGACCCCA	898	
Db	1081	ATT	CTCAATGCAACCGAAAAATGCTCTCTTCG	GAATATCTTCTCCTTAATTCGACCCCA	1140	
Qy	899	GT	CGTGTAAACGATGATGTAGAAAGATCTTTAA	AAACAAAACCTGATGCGGAAAAAATGAC	958	

Db	1141	 GTCGTGAACGATGATGATGACGAAAGATCTTAAAAAATACTGATGCGGGAATAATGAC	1200
Qy	959	AATCAGACAGCCGCGCGCCGCAAGAAATTCGACAAATTAATCGAAAAATTTGAAGCTGAATG	1018
Db	1201	AATCAGACAGCCGCGCGCCGCAAGAAATTCGACAAATTAATCGAAAAATTTGAAGCTGAATG	1260
Qy	1019	CGCAGAAAGTTTGGGATAACGAAAT-----	1042
Db	1261	CGCAGAAAGTTTGGGATAACGAAATGTTAGTTTAAATTAATTCAAACAATTAATATACAAT	1320
Qy	1043	-----GTGAGATTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGGAAA	1092
Db	1321	TGATTTTCAGGTGAGATTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGGAAA	1380
Qy	1093	ACTCTCTGTTTATAAAGTCACTGGTAAATTCGACAGAGAGAAATGCAAAAAAGTACG	1152
Db	1381	ACTCTCTGTTTATAAAGTCACTGGTAAATTCGACAGAGAGAAATGCAAAAAAGTACG	1440
Qy	1153	ATACTCAATTTGTTCAAAATCTATGAGGAAAAAATAAGTTTCATTTAGTTTTCGCCACTAC	1212
Db	1441	ATACTCAATTTGTTCAAAATCTATGAGGAAAAAATAAGTTTCATTTAGTTTTCGCCACTAC	1500
Qy	1213	CATGAGTCAAAAGTTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATGGAAATCTTTGAAG	1272
Db	1501	CATGAGTCAAAAGTTTAAAGTGGAGCAAAAGAAATACGCTGTACCAATGGAAATCTTTGAAG	1560
Qy	1273	TTCAATGAGAGCCCAAAAGATACAGAAATCGAATTTGATCTGGTGATGCAAGACAAGTTTC	1332
Db	1561	TTCAATGAGAGCCCAAAAGATACAGAAATCGAATTTGATCTGGTGATGCAAGACAAGTTTC	1620
Qy	1333	TAAAGCGAGCTACACGAAAAACCTCACGACTACAAAGAAAAATACCCATAAAATGCTGAAAG	1392
Db	1621	TAAAGCGAGCTACACGAAAAACCTCACGACTACAAAGAAAAATACCCATAAAATGCTGAAAG	1680
Qy	1393	AAATGGAATTTCTTCTGTAAGAGCTAAATTTGTTGAAAGATTTGGAATTAATGCTCCAAAC	1452
Db	1681	AAATGGAATTTCTTCTGTAAGAGCTAAATTTGTTGAAAGATTTGGAATTAATGCTCCAAAC	1740
Qy	1453	TTCAATGATGCAATGTGCGAGAAAGTTTGAAGAGCCAAATGCTGTGCAATAGTGTAA	1512
Db	1741	TTCAATGATGCAATGTGCGAGAAAGTTTGAAGAGCCAAATGCTGTGCAATAGTGTAA	1800
Qy	1513	ATGAAACAAATTTAAATGACACGAGTGAATTCGTGCAATTTCAAGAAAAACAATTCGAATGTG	1572
Db	1801	ATGAAACAAATTTAAATGACACGAGTGAATTCGTGCAATTTCAAGAAAAACAATTCGAATGTG	1860
Qy	1573	TTCCGAAAAAGAACTTTGCTGTGCTGTTTTTGTAGTCAACGAAACAGCGGGAATCCAT	1632
Db	1861	TTCCGAAAAAGAACTTTGCTGTGCTGTTTTTGTAGTCAACGAAACAGCGGGAATCCAT	1920
Qy	1633	GCTTAGAAGAGACGAGCTGT-----	1654
Db	1921	GCTTAGAAGAGACGAGCTGTGTTGTTGTTTCTACGTAGATTAATTCGGAATATTTTC	1980
Qy	1655	--TAAAGTTCTACACCGAACTAAATTTGGTGGTTGCAAGTTCGTCGGAATACGAATGCTGCC	1712
Db	1981	AGTAAGTTCTACACCGAACTAAATTTGGTGGTTGCAAGTTCGTCGGAATACGAATGCTGCC	2040
Qy	1713	AATGAAAAACAGAGGAGCGCAATCTATTATGTACGACGCAAGAAAAATGAATATGCC---	1769
Db	2041	AATGAAAAACAGAGGAGCGCAATCTATTATGTACGACGCAAGAAAAATGAATATGCCGTA	2100
Qy	1770	-----TTCTCAAAAAATTTGTACAC	1788
Db	2101	AGTTTTCAGAAAAATGGAAGTTTTTAAATATCAATTAATTTACAGTTTCTCAAAAAATTTGTACAC	2160
Qy	1789	TAAATACCGGAATCGGTAGATTTGAAATAGCCCAACAGAGCGGAAGATATGTTTGAAC	1848
Db	2161	TAAATACCGGAATCGGTAGATTTGAAATAGCCCAACAGAGCGGAAGATATGTTTGAAC	2220
Qy	1849	GTCTTCCGATTAAGAAACAAAAAGTCTTAATGTTTCAATTAATTTCCAAACGACAATGTA	1908

Db 2221 GTCTTCCGATAAAGAACAAAAGTCTTAATGTTCAATTATCATTTCCAAACGACAACCTGA 2280
 QY 1909 ATGCTTACCGGTTTTGTGAACAATTAATGCGATCAACCAATCGGTGTAGCTAATCAAGCATTA 1968
 Db 2281 ATGCTTACCGGTTTTGTGAACAATTAATGCGATCAACCAATCGGTGTAGCTAATCAAGCATTA 2340
 QY 1969 TTACTTCTCAACACAGTCAACAAGCTTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2028
 Db 2341 TTACTTCTCAACACAGTCAACAAGCTTTTGGCATCACTAAGGCACGAGAAAGGATCAAAAC 2400
 QY 2029 GAATTTTCTATCAAAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2088
 Db 2401 GAATTTTCTATCAAAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2460
 QY 2089 TTGACTGGTTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 2148
 Db 2461 TTGACTGGTTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 2520
 QY 2149 CATTAACTATGATGTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2208
 Db 2521 CATTAACTATGATGTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 2580
 QY 2209 ATTCTATAGCGGTGTAGTAGCGAGTATCAATCCAGGTGGAATTCATCTATCGAAATATGA 2268
 Db 2581 ATTCTATAGCGGTGTAGTAGCGAGTATCAATCCAGGTGGAATTCATCTATCGAAATATGA 2640
 QY 2269 TTGTGACTCAAGAAGATGTGTCCTCGGTGAGGTGCGATGCGTGTATGAGCGGAAAGAA 2328
 Db 2641 TTGTGACTCAAGAAGATGTGTCCTCGGTGAGGTGCGATGCGTGTATGAGCGGAAAGAA 2700
 QY 2329 CAGATATTTTGGAGCAAAAGTTTCGTGAATTCGTCAAGAAATTCGTCAAGAAATTCGTCAAGAAATTCGT 2378
 Db 2701 CAGATATTTTGGAGCAAAAGTTTCGTGAATTCGTCAAGAAATTCGTCAAGAAATTCGTCAAGAAATTCGT 2760
 QY 2379 -----AACACGAC 2387
 Db 2761 TTGAGTATTTAAAGATCTCGGATTTTAAATTTTGTAACTTTTCAGAAACACGAC 2820
 QY 2388 AATCGAGCACCGCGCATATTTAGTCTATCGAGACGGAGTTAGCGATTCGAGATGCTTA 2447
 Db 2821 AATCGAGCACCGCGCATATTTAGTCTATCGAGACGGAGTTAGCGATTCGAGATGCTTA 2880
 QY 2448 CGTGTGTAGTCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAATTCATGTCGAA 2507
 Db 2881 CGTGTGTAGTCATGATGAGCTTCGATCTTTTAAAGCGAAGTAAACAATTCATGTCGAA 2940
 QY 2508 CGGGATGGAGAAGATCCAGACCGAAGTACAGTTCATTTGATTCAGAAAGACACAAAT 2567
 Db 2941 CGGGATGGAGAAGATCCAGACCGAAGTACAGTTCATTTGATTCAGAAAGACACAAAT 3000
 QY 2568 ACACGATTCGTTTGAAGAATGGAAAGATAGACCGAGTGTCAATTAAGATCTTACTCT 2627
 Db 3001 ACACGATTCGTTTGAAGAATGGAAAGATAGACCGAGTGTCAATTAAGATCTTACTCT 3060
 QY 2628 GCTGAAACAGATGTCGCTGTGCTGTTTAAACATGGAGAGGATAGAAAGAAC 2687
 Db 3061 GCTGAAACAGATGTCGCTGTGCTGTTTAAACATGGAGAGGATAGAAAGAAC 3120
 QY 2688 AAAGAACTCGAATTTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCG 2747
 Db 3121 AAAGAACTCGAATTTGTGAACCCATCATCCGGAACAACTGTGGATAAACTTATCGTTTCG 3180
 QY 2748 AAATACAAATTCGATTTTCTTGGCATCTCATGTCATGTCCTTGGTACATCTCGTCCA 2807
 Db 3181 AAATACAAATTCGATTTTCTTGGCATCTCATGTCATGTCCTTGGTACATCTCGTCCA 3240
 QY 2808 GGACATTTACACTGTTATGATGACGATAAAGGAATGAGCAAGATCAAGTCTAT ----- 2861
 Db 3241 GGACATTTACACTGTTATGATGACGATAAAGGAATGAGCAAGATCAAGTCTATGTAAGC 3300
 QY 2862 ----- 2861
 Db 3301 GTTTTGAATAGCAGTTAGCGATTTTGTAGGATTTTGTAAATCCGCATATAGTTATTATATAAAA 3360

QY 2862 -----RAAATGACCTAGGACTTGCCTTTCTCTCTAGATGTGAAAAACCCAT 2911
 Db 3361 AATGTTTCAGAAAAATGACCTAGGACTTGCCTTTCTCTCTAGATGTGAAAAACCCAT 3420
 QY 2912 CTCGTGTGCTGTTCGCGTTTCATTATGCTCATTTATCATGTGAAAAAGCGAAAGAGCTTTTA 2971
 Db 3421 CTCGTGTGCTGTTCGCGTTTCATTATGCTCATTTATCATGTGAAAAAGCGAAAGAGCTTTTA 3480
 QY 2972 TCGAACTTTACAAGGAACATTATCATCGGTGACTATGTCACAGCCACGGAATTCGACACGAAAT 3031
 Db 3481 TCGAACTTTACAAGGAACATTATCATCGGTGACTATGTCACAGCCACGGAATTCGACACGAAAT 3540
 QY 3032 GGAACATTTTCTCCAACTAAGCTGAAGTACCCTGGAATGCTGTCGATACATTTTGC 3091
 Db 3541 GGAACATTTTCTCCAACTAAGCTGAAGTACCCTGGAATGCTGTCGATACATTTTGC 3600
 QY 3092 AAAAGTGTGCGCGGTTTCAATCAAAATTTTCAATTTGTAGATATTTGACTTTACTTTTTTTT 3151
 Db 3601 AAAAGTGTGCGCGGTTTCAATCAAAATTTTCAATTTGTAGATATTTGACTTTACTTTTTTTT 3660
 QY 3152 AAAGCCCGGTTTCAAAATTTCAATCCATGACTAAAGTGTGTCATTAATTTACTTTGAAATTT 3210
 Db 3661 AAAGCCCGGTTTCAAAATTTCAATCCATGACTAAAGTGTGTCATTAATTTACTTTGAAATTT 3719

RESULT 7

US-10-174-363-3
 ; Sequence 3, Application US/10174363
 ; Publication No. US20030077623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Butler, Karlene H.
 ; APPLICANT: Cahoon, Rebecca E
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Rafalski, Antoni J.
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptiona
 ; TITLE OF INVENTION: Gene Silencing
 ; FILE REFERENCE: BB1454 US NA
 ; CURRENT APPLICATION NUMBER: US/10/174,363
 ; PRIOR FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: 60/298,973
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 3
 ; LENGTH: 2827
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-174-363-3

Query Match 2.0%; Score 63.8; DB 6; Length 2827;
 Best Local Similarity 52.0%; Pred. No. 0.00026;
 Matches 143; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
 QY 2674 ATATCAAGAAGCAAGAAAGAACTGGAAATTTGCAACCCATCATCCGGAACAACTGTGGATA 2733
 Db 2028 ACAGGACAGAGAGCAGTACAGATCGAGTGGGAATATATTGCTGGGACTGTGTTGATA 2087
 QY 2734 AACTTATCGTTTCGAAATACAAATTCGATTTTTTCTTGGCATCTCATCATGTTGTCCTTG 2793
 Db 2088 CCAAAATCTGCCATCCACAGAAATTTGATTTTTTATCTCTCGAGCCATGCTGGCATCCAG 2147
 QY 2794 GTACATCTCGTCAGACATTTACACTGTATGTATGACGATAAAGGAATGAGCCAGATG 2853
 Db 2148 GTACTAGTCGGCAGCTCATTTATCATGTCCTGCTGGGATGAAAAACAACCTTCACACCTGATG 2207
 QY 2854 AAGTCTATAAAATGACCTACGAGCTTGTCTTCTCTGCTAGATGTTCGAAAAACCATCT 2913
 Db 2208 GAATTCAGTCTCTGACAAACAACCTTTGTTATACATATGCCAGGTGTACACGCTCAGTAT 2267
 QY 2914 CGTTCGCTGTTCCGGTTCAATTATGCTCATTTATCA 2948

Db 2268 CAGTTGTTCTCCAGCATATATATGACACATTTAGCA 2302

RESULT 8

US-11-093-888-3

; Sequence 3, Application US/11093888

; Publication No. US20050204427A1

; GENERAL INFORMATION:

; APPLICANT: Butler, Karlene H.

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Harvelli, Leslie T.

; APPLICANT: Rafalecki, Antoni J.

; APPLICANT: Sakai, Hajime

; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptional

; TITLE OF INVENTION: Gene Silencing

; FILE REFERENCE: BBI454 US NA

; CURRENT APPLICATION NUMBER: US/11/093,888

; CURRENT FILING DATE: 2005-03-30

; PRIOR APPLICATION NUMBER: US/10/174,363

; PRIOR FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: 60/298,973

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 3

; LENGTH: 2827

; TYPE: DNA

; ORGANISM: Glycine max

; US-11-093-888-3

Query Match 2.0%; Score 63.8; DB 13; Length 2827;

Best Local Similarity 52.0%; Pred. No. 0.00026;

Matches 143; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 2674 ATATGAAGAAGCAAGCAAACTGGAAATGTGAACCCATCATCGGAAACAACTGTGGATA 2733

Db 2028 ACAGGGACAGACAGCAGTACAGATCGGAGTGGGAATATATGCTGGGACATGTTGGATA 2087

QY 2734 AACTATCGTTTGAATACAAATTCGATTTTCTGTTGGCATCTCATCATGTTGCTCTTG 2793

Db 2088 CCAAAATTCGCCATCCCAAGAAATTTGATTTTATCTCTGCAGCCATCTGGCATCCAGG 2147

QY 2794 GTACATCTCTCCAGGACATTACACTGTTATGTATGACGATAAGGAATGAGCCAGATG 2853

Db 2148 GTHACTAGTCGGCAGCTCATATCATGCTCTGTGGGATGAACAACTTACACCTGATG 2207

QY 2854 AAGTCTATAAATGACCTACGAGCTTGCTTTCTCTCTGCTAGATGCGAAACCCATCT 2913

Db 2208 GAATTCAGTCTCTGACAAACAACTTTGTTATACATATGCCAGTGTACAGCTCAGTAT 2267

QY 2914 CGTTGCTGTTCCGGTTTCATTATGCTCATTTATCA 2948

Db 2268 CAGTTGTTCTCCAGCATATATATGACACATTTAGCA 2302

RESULT 9

US-09-918-995-10623

; Sequence 10623, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10623

; LENGTH: 490

; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10623

Query Match 1.9%; Score 62.4; DB 3; Length 490;

Best Local Similarity 51.8%; Pred. No. 0.00021;

Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2698 GAATTGTGAACCCCATCATCGGAACTGTGGATAAACTTATCGTTTCGAAATACAAAT 2757

Db 215 GAATGTCATATCCAGCTGGAAACAGTTGTATACAGATTACACACCCATATGAGT 274

QY 2758 TCGATTTTCTTGGCATCTCATCATGTTGCTTGGTGTACATCTCGTCCAGGACATTACA 2817

Db 275 TCGATTTTACCTCTGTAGCCATCTGGAATACAGGGTACCAGTCGTCCTTTCACACTATC 334

QY 2818 CTGTTATGTATGAGATAAAGGAATGAGCAAGATGAAGTCTATAAAATGACCTACGGAC 2877

Db 335 ATGTTTATGGGATGATAACTGCTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 394

QY 2878 TTGCTTTTCTCTGCTAGATGCGAAACCCATCTCGTTGCCCTGTTCCGGTTTCATTATG 2937

Db 395 TCTGCCACACTTAGCTAGCTGTACAGATCTGTTTCTATACCTGCACCCAGCGTATTATG 454

QY 2938 CTCATTTTATCATGTGAAAAAGCGAAAGAGCTT 2969

Db 455 CTCACCTGTAGCATTTAGAGCCAGATATCAT 486

RESULT 10

US-10-779-543-3372

; Sequence 3372, Application US/10779543

; Publication No. US20050227917A1

; GENERAL INFORMATION:

; APPLICANT: Williams et al

; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED

; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II

; FILE REFERENCE: 2300-21302

; CURRENT APPLICATION NUMBER: US/10/779,543

; PRIOR FILING DATE: 2004-02-12

; PRIOR APPLICATION NUMBER: 10/076,555

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: 09/217,471

; PRIOR FILING DATE: 1998-12-21

; PRIOR APPLICATION NUMBER: 60/068,755

; PRIOR FILING DATE: 1997-12-23

; PRIOR APPLICATION NUMBER: 60/080,664

; PRIOR FILING DATE: 1998-04-03

; PRIOR APPLICATION NUMBER: 60/105,234

; PRIOR FILING DATE: 1998-10-21

; PRIOR APPLICATION NUMBER: 09/297,648

; PRIOR FILING DATE: 2000-04-10

; PRIOR APPLICATION NUMBER: PCT/US99/01619

; PRIOR FILING DATE: 1999-01-28

; PRIOR APPLICATION NUMBER: 60/072,910

; PRIOR FILING DATE: 1998-01-28

; PRIOR APPLICATION NUMBER: 60/075,954

; PRIOR FILING DATE: 1998-02-24

; PRIOR APPLICATION NUMBER: 60/080,114

; PRIOR FILING DATE: 1998-03-31

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 23767

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3372

; LENGTH: 734

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 5, 6, 7, 8, 9, 10, 26, 639


```

; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 3578
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-37

Query Match      1.94; Score 62.4; DB 10; Length 3578;
Best Local Similarity 51.84; Prsd. No. 0.00066;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy      2698  GAATTGTGAACCCATCATCCGGAAACAACCTGTGGATAAACTTATCGTTTCGAATACAAAT 2757
          |||||
Db      2436  GNAATGGCAATATCCAGCTGGAAACAACAGTTGATACAGACATTACACACCCATATGAGT 2495

Qy      2758  TCGATTTTTTCTTGGGCATCTCATCATGGTGTCTTGGTACATCTCGTCCAGGACATTACA 2817
          |||||
Db      2496  TCGATTTTTTACCTCTGTAGCCATCTGTGGAAATCAGGGTACCAAGTCGCTTTCACACTATC 2555

Qy      2818  CTGTTATGTATGACGATAAGGAATCAGCCAAAGTGAAGTCTATAAAATGACCTACGGAC 2877
          |||||
Db      2556  ATGTTTTATGGGATGATACTGCTTTTACTGCAGATGAACTTCAGCTGCTAACTTACCAGC 2615

Qy      2878  TTGCTTTTCTCTGCTAGATGTCGAAAACCCCATCTCGTTGCTGTTCGGTTTCATTATG 2937
          |||||
Db      2616  TCTGCCACACTTACGTACGCTGTACACGATCTGTTTCTATACCTCGCACCAAGCGTATTATG 2675

Qy      2938  CTCATTATCATGTGAANAAGCGAAGAGCTT 2969
          |||||
Db      2676  CTCACCTGGTAGCATTTAGAGCCAGATATCAT 2707

RESULT 15
RES-10-174-163-5

```

```

; Sequence 5, Application US/10174363
; Publication No. US2003007623A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptional Regulation Of Gene Expression
; TITLE OF INVENTION: Gene Silencing
; FILE REFERENCE: BB1454 US NA
; CURRENT APPLICATION NUMBER: US/10/174,363
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,973
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Glycine max
US-10-174-363-5

Query Match          1.9%;   Score 60;   DB 6;   Length 1501;
Best Local Similarity 51.9%;   Pred. No. 0.0016;
Matches 135; Conservative      0; Mismatches 125; Indels      0; Gaps      0;

Qy    2700 ATTGTGAACCCATCATCCGGAAACAATTGTGGATAAACCTTATCGTTTCGAAATACAAATTC 2759
       |||||
Db    851 AGTGGAATATCTTACCCTGGTACTGTGGTGGATTCTTAAGATCTGTCTACCGGAATTC 910
       |||||

Qy    2760 GAATTTTTCTTGCATCTCATCATGTGTGCCCTTGGTAGATCTCGTCCAGGACATTACACT 2819
       |||||
Db    911 GACTTCTATTATTATGCACTCATCCGGGAATTCAGGGGTACAAGTAGACCAAGCTCATTATCAT 970
       |||||

Ov    2820 GTTATGTATGACGATAAAGCAATGAGCCAGATGAAGTCTATAAAAATGACTTACCGGACTT 2879

```

Db	971	GTCTCTGGGACGAGAACAAATTTCAC	TGCTGATGAGATCCATCTCTGACCACTTG	1030
Qy	2880	GCTTTTCTCTGCTAGATGTCGAAAA	CCCATCTCGTTGCCTGTTCCGGTTCATTATGCT	2939
Db	1031	TGCTACACCTATGCAAGATGTACACG	ATCAGTTTCTGTAGTGCCTCCTGCGTACTATGCT	1090
Qy	2940	CATTTATCATGTGAAAAAGC		2959
Db	1091	CATTGGCAGCTTACAGAGC		1110

Search completed: July 5, 2006, 23:14:00
Job time : 3768 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 22:10:12 ; Search time 385 Seconds
(without alignments)
10897.477 Million cell updates/sec

Title: US-10-645-746-2
Perfect score: 3227
Sequence: 1 cagccacaagtgatgaac.....tttaaaaaaaaaaaaaa 3227

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:
1: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:
2: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:
4: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:
5: /EMC Cellerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:
6: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:
7: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:
8: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:
9: /EMC Cellerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.4	1.9	3320	8	US-11-266-748A-31508
2	62.4	1.9	3320	8	US-11-266-748A-31508
3	61	1.9	1196	6	US-10-449-902-3167
4	61	1.9	1196	6	US-10-449-902-3167
5	61	1.9	1196	6	US-10-449-902-3167
6	56.4	1.7	4758	8	US-11-266-748A-30266
7	54.6	1.7	1211	6	US-10-449-902-26400
8	54	1.7	1094	7	US-11-218-305-21995
9	53.6	1.7	892	7	US-11-218-305-19386
10	52.8	1.6	3350	7	US-11-218-305-19388
11	52	1.6	1710	7	US-11-218-305-19385
12	51	1.6	3474	6	US-10-953-349-4063
13	50.2	1.6	3422	6	US-10-449-902-27837
14	46.6	1.4	964	7	US-11-218-305-22488
15	46.6	1.4	965	7	US-11-218-305-22488
16	46.6	1.4	2001	7	US-11-218-305-22483
17	46.2	1.4	1816	6	US-10-953-349-19397
18	45.8	1.4	1469	7	US-11-218-305-22082
19	45.6	1.4	3394	8	US-11-266-748A-26618
20	45.6	1.4	3394	8	US-11-266-748A-26618
21	45.2	1.4	2016	7	US-11-218-305-22484
22	45.2	1.4	3463	6	US-10-449-902-26806
23	45.2	1.4	3746	7	US-11-218-305-22486
24	45	1.4	997	8	US-11-216-545-6851

25	44.6	1.4	1793	8	US-11-266-748A-25558	Sequence 25558, A
26	44.4	1.4	3397	6	US-10-517-441-378	Sequence 378, App
27	44.4	1.4	3397	6	US-10-517-441-378	Sequence 378, App
28	44.2	1.4	1000	8	US-11-266-748A-394325	Sequence 394325, App
29	44.2	1.4	1000	8	US-11-266-748A-465371	Sequence 465371, A
30	44.2	1.4	7478	8	US-11-266-748A-56189	Sequence 56189, A
31	43.6	1.4	903	7	US-11-218-305-19468	Sequence 19468, A
32	43.6	1.4	2095	7	US-11-218-305-1556	Sequence 1556, App
33	43.2	1.3	1555	8	US-11-216-545-7032	Sequence 7032, App
34	43	1.3	1934	8	US-11-216-545-6852	Sequence 6852, App
35	42.8	1.3	2773	7	US-11-218-305-19467	Sequence 19467, A
36	42.4	1.3	2503	6	US-10-449-902-19320	Sequence 19320, A
37	41.8	1.3	1000	8	US-11-266-748A-290582	Sequence 290582, A
38	41.8	1.3	1000	8	US-11-266-748A-342011	Sequence 342011, A
39	41.8	1.3	1000	8	US-11-266-748A-401890	Sequence 401890, A
40	41.8	1.3	1000	8	US-11-266-748A-472936	Sequence 472936, A
41	41.8	1.3	294540	8	US-11-266-748A-23953	Sequence 23953, A
42	40.8	1.3	951	8	US-11-266-748A-48160	Sequence 48160, A
43	40.8	1.3	1298	8	US-11-216-545-2984	Sequence 2984, App
44	40.6	1.3	713	8	US-11-266-748A-211284	Sequence 211284, App
45	40.6	1.3	713	8	US-11-266-748A-234991	Sequence 234991, A

ALIGNMENTS

RESULT 1
US-11-266-748A-31508
; Sequence 31508, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31508
; LENGTH: 3320
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-31508

Query Match 1.9%; Score 62.4; DB 8; Length 3320;
Best Local Similarity 51.8%; Pred. No. 3.4e-05;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 2698 GAATTGTGAACCCATCATCGGAACAACTGTGGTAAACTTATCGTTTTCGAATAACAAT 2757
Db 2178 GAAGTGGCAATATCCAGCTGGGAACAACAGTTGATACAGACATTACACACCATATGAGT 2237
QY 2758 TCGATTTCCTTGGCATCTCATCGTGGTGTCTCTGTGTACATCTCGTCCAGGACATTACA 2817

Db 2238 TCGATTTTACCTCTGTAGCCATGCTGGAATACAGGGTACCAGTCTGCTTCCACACTATC 2297
Qy 2818 CTGTTATGATGACGATAAGGAATGAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
Db 2298 ATGTTTTATGGGATGATAACTGCTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 2357
Qy 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGTTCCGGTTCATTATG 2937
Db 2358 TCTGCCACACTTACGTACGCTGTACACGATCTGTTTCTATACCTGCACCGGCTATTATG 2417
Qy 2938 CTCATTATCATGTGAAAAGCGAAAGAGCTT 2969
Db 2418 CTCACCTGCTAGCATTTAGAGCCAGATATCAT 2449

RESULT 2

US-11-266-748A-56999
; Sequence 56999, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56999
; LENGTH: 3320
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-56999

Query Match 1.9%; Score 62.4; DB 8; Length 3320;
Best Local Similarity 51.8%; Pred. No. 3.4e-05;
Matches 141; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
Qy 2698 GAATGTGAACCCATCATCGGNACAACTGTGGATAAACTTATCGTTTCCAAATACAAT 2757
Db 2178 GAAGTGGCAATATCCAGCTGGAAACAACTGTTGATACAGACATTACACACCCATATGAT 2237
Qy 2758 TCGATTTTTCTTGGCATCTCATCATGTTGTCCTTGGTATACATCTCGTCCAGGACATTACA 2817
Db 2238 TCGATTTTTCTTGGCATCTCATCATGTTGTCCTTGGTATACATCTCGTCCAGGACATTACA 2817
Qy 2818 CTGTTATGATGACGATAAGGAATGAGCCAGATGAAGTCTATAAATGACCTACGGAC 2877
Db 2298 ATGTTTTATGGGATGATAACTGCTTTACTGCAGATGAACCTTCAGCTGCTAACTTACCAGC 2357
Qy 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGTTCCGGTTCATTATG 2937
Db 2358 TCTGCCACACTTACGTACGCTGTACACGATCTGTTTCTATACCTGCACCGGCTATTATG 2417
Qy 2938 CTCATTATCATGTGAAAAGCGAAAGAGCTT 2969

Db 2418 CTCACCTGCTAGCATTTAGAGCCAGATATCAT 2449

RESULT 3

US-10-449-902-3167
; Sequence 3167, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3167
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK060907
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-3167

Query Match 1.9%; Score 61; DB 6; Length 1196;

Best Local Similarity 53.1%; Pred. No. 4.7e-05;
Matches 130; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
Qy 2716 CGGAAACAACTGTGGATAAACTTATCGTTTCCGAAATACAAAATTCGATTTTTCTTGGCAT 2775
Db 517 CAGGTACTGTTGTGGACAAACAAAGTGTGCCATCCAGGNAATTATGACTTCTACATGTG 576
Qy 2776 CTCATCATGTGTCTCTTGGTACATCTCGTCCAGGACATTPACACTGTTATGTATGACGATA 2835
Db 577 CTCATGCTGGAATGATTGGAACGAGCGCAACACATTTATCATGTCTTGCACGATGAGA 636
Qy 2836 AAGGAATGAGCCAGATGAAGTCTATAAAATGACCTACGGACTTGTCTTCTCTCGCTA 2895
Db 637 TAGGTTTCTCCCTCATGATCTGCGAGGAGTAGTGCACTCTCTTTAIGTATCAGA 696
Qy 2896 GATGTCGAAAACCCATCTCGTTGCTTCCGTTTCATTATGCTCATTTTATCATGTGAAA 2955
Db 697 GNAGCACACACAGCCATATCAGTTGTTGCTCCGATTTGCTATGCCCATCTTGTCTGCTC 756
Qy 2956 AAGCG 2960
Db 757 AGGTG 761

RESULT 4

US-10-449-902-3806
; Sequence 3806, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870


```
QY      2939 TC 2940
Db      2614 CC 2615

RESULT 7
US-10-449-902-26400
; Sequence 26400, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; FILE OF INVENTION: FULL-LENGTH PLANT cDNA AND USSES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26400
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK101842
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-26400

Query Match      1.7%; Score 54.6; DB 6; Length 1211;
Best Local Similarity 49.1%; Pred. No. 0.0018;
Matches 144; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY      2656 TTAAACAATGGGAGGATATGAAGAAGCAAGCAAACTGGAATTTGGAACCCATCAT 2715
Db      493 TTGCTCAAAGAATCAACCACAAACATTTTCCAGAAAATACTCCAGACAAAGTTCCCTC 552

QY      2716 CGGGAACAATCTGGATAAATTTATCTGTTTCGAAATACAAATTCGATTTTCTTGGCAT 2775
Db      553 CAGGGACTGTGTAGATTTCTGGTATTTGTTTCATCCAGACAGATGATTTCTACATGATG 612

QY      2776 CTCATCATGTGTCTTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACGATA 2835
Db      613 CTCATGCTGGACCTATAGGTACCTCAAGGCCACCATTTACCATTGCTTGTCTCGATGAGA 672

QY      2836 AAGGAATGACCCAGATGAAGTCTATAAAATGACCTACGGACTTGCTTTTCTCTCTGCTA 2895
Db      673 TTGGTTTCTTGCCCGATGACGTGCAGAAACTAGTTTTTGTGCGTTTCTTATGTGTATCAGA 732

QY      2896 GATGTGAAACCCATCTCGTTGCGCTGTTCCGGTTTCATTATGCTCATTTATCA 2948
Db      733 GGAGCACCACTGCAATATCTGTTGGCACCTATCTGTTATGCTCATCTCGCA 785

RESULT 8
US-11-218-305-21995/c
; Sequence 21995, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19386
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-21995

Query Match      1.7%; Score 53.6; DB 7; Length 892;
Best Local Similarity 51.7%; Pred. No. 0.0027;
Matches 122; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY      2713 CATCCGGAACAACCTGTGGATAAATTTATGTTTCGAAATACAAATTCGATTTTCTTGG 2772
Db      238 CAGCCGGCACCTGTGTTGACAAACAAGTCTGTCTATCCAAAGAACTTCGATTTCTACATGT 297

QY      2773 CATCTCATCATGTGTCTTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACG 2832
Db      298 GTTCATGCTGGAATGATTGGGACTACACAGGCCAACTCACTATCATCATCTCTGATGACG 357

QY      2833 ATAAAGGAATGAGCCCAAGATGAAGTCTTATAAAATGACCTTACGAGACTGCTTTCTCTG 2892
```


Db 358 AGATAGGCTTCAATCTGATGACCTGCAGGAGCTGGTGCACTCGCTCTCTTATGTGTACC 417
QY 2893 CTGATGTCGAACCAACCATCTGTTGCTGTTCCGGTTCAATTATGCTCATTTATCA 2948
Db 418 AAAGGAGCAACAAGCCATATCAGTTGTTGCTCCCATCTGCTAGCAGCACCTTTGCA 473

RESULT 10

US-11-218-305-19388
; Sequence 19388, Application US/11218305
; Publication No. US2006014195A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19388
; LENGTH: 3350
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (738)..(742)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908)..(930)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3235)..(3236)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-19388

Query Match 1.6%; Score 52.8; DB 7; Length 3350;
Best Local Similarity 51.2%; Pred. No. 0.0077;
Matches 123; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 2713 CATCCGGAACAACCTGTGGATAAATCTATCGTTTCGAAATACAAATTCGATTTTCTTGG 2772
Db 2525 CACCAGAACTGGTGGACACAAGAAGTCTGCCATCCAAGAACTTCGATTTCTACATGT 2584
QY 2773 CATCTCATCATGTCCTTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACG 2832
Db 2585 GTGCGCATGCTGGAATGCGGACTACGAGGCCAACTCACTACCACATCCTGCATGATG 2644
QY 2833 ATAAAGGAATGAGCCAAAGATGAAGTCTATAAAATACCTACGGAATGCTTTTCTCTG 2892
Db 2645 AGATAGGCTTTCAGTCTCCTGATGATCTGCAGGAGCTGGTGCAATTCGCTCTCTTATGTACC 2704
QY 2893 CTAGATGTCGAACCACTCGTTGCCCTTCCGTTTCATTTATGCTCATTTATCATGTG 2952
Db 2705 AAAGGAGCAACAAGCCATATCAGTCTGTTGCTCCCATCTGCTACGCAACATCTGGCAGCTG 2764

RESULT 11

US-11-218-305-19385
; Sequence 19385, Application US/11218305
; Publication No. US2006014195A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng

; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19385
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-19385

Query Match 1.6%; Score 52; DB 7; Length 1710;
Best Local Similarity 51.3%; Pred. No. 0.009;
Matches 121; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 2713 CATCCGGAACAACCTGTGGATAAATCTATCGTTTCGAAATACAAATTCGATTTTCTTGG 2772
Db 1040 CAGCCGGCACTGTTGTTGACACAAGAAGTCTGTATCCAGGAATTCGATTTCTACATGT 1099
QY 2773 CATCTCATCATGTCCTTGGTACATCTCGTCCAGGACATTACACTGTTATGTATGACG 2832
Db 1100 GTTCACATGCTGGAATGATTGGAACCTACCAGGCCAACTCACTATCATCCTCGCATGACG 1159
QY 2833 ATAAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACGGAATTCGTTTCTCTCTG 2892
Db 1160 AGTAGGCTTCAATCTCTGATGACCTCGCAGGAGCTGGTGCACTCGCTCTCTTACGTGTACC 1219
QY 2893 CTAGATGTCGAAAACCCATCTCGTTGCCCTGTTCCGGTTTCATTTATGCTCATTTATCA 2948
Db 1220 AAAGGAGCAACAAGCCATATCAGTTGTTGCTCCCATCTGCTACGCCACCTTGCA 1275

RESULT 12

US-10-953-349-4063
; Sequence 4063, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4063
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-4063

Query Match 1.6%; Score 51; DB 6; Length 3474;
Best Local Similarity 50.2%; Pred. No. 0.022;
Matches 126; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 2698 GAATGTGAAACCATCATCCGGAACAACTGTGTGATAACTTATCGTTTCGAAATACAAAT 2757
Db 2861 GAAGTGGGAATATTTTACCTGGCACTGTTGTGGACTCTAAAATCTGCCACCTTACAGAGT 2920
QY 2758 TCGATTTTCTTGGCATCTCATCATGTTGTTGTTGATATCTCTTGGTATCATCTCGTCCAGGACATTACA 2817
Db 2921 TTGACTTTTACCTCTGTAGTCACTGTGTTTACGGGCACTTCTCGACCTGTCTCATACC 2980
QY 2818 CTGTTATGTATGACGATAAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 2981 ACGTTCTTTGGGATGAGAACAACTTTACTGCAGATGGACTTCAATCTCTGACCAATACT 3040
QY 2878 TTGCTTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGCTGTTCCGGTTCATTATG 2937

Db 3041 TATGTTACAGTATGCAAGTGCACACGCTCAGTTTCATTGTTCCCTCGCATATTATG 3100
QY 2938 CTCATTATCA 2948
| | | | |
Db 3101 CACATCTAGCA 3111

RESULT 13

US-10-449-902-27837
; Sequence 27837, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27837
; LENGTH: 3432
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK103279
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-27837

Query Match 1.6%; Score 50.2; DB 6; Length 3432;
Best Local Similarity 50.2%; Pred. No. 0.034;
Matches 124; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 2698 GAATTGTGAACCATCATCCGGAACAACTGTGGATAACTTATCGTTTGGAAATACAAAT 2757
Db 2743 GGAGCGGGAACATACTCCCTGGTACGTTGTAGATTCAAGATTCCGTCATCCAACTGAGT 2802
QY 2758 TCGATTTTCTTGGCATCTCATCATGCTCTCTTGTACATCTCGTCCAGGACATTACA 2817
Db 2803 TTGACTTCTACCTGTGTAGCCATGCTGGCATTAAGGGTACTAGTCGTCCAGCTCATATC 2862
QY 2818 CTGTTATGTATGACGATAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACGGAC 2877
Db 2863 ATGCTCTGTGGATGAACAACACTTCACAGCTGATGCATTGCAGATTCTTACCAACAAC 2922
QY 2878 TTGCTTTCTCTGCTAGATGTCGAAACCCATCTCGTTGCTGTTCCGGTTTCATTATG 2937
Db 2923 TTTGCTACACCTATGCAAGGTGCACTCGCTCTGTATCAATTGTTCCACCTGCTTATTATG 2982
QY 2938 CTCATTT 2944
Db 2983 CTCATCT 2989

RESULT 14

US-11-218-305-22488
; Sequence 22488, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305

; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22488
; LENGTH: 964
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: [941]..(942)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: [945]..(950)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: [952]..(953)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: [963]..(964)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-22488

Query Match 1.4%; Score 46.6; DB 7; Length 964;
Best Local Similarity 49.0%; Pred. No. 0.15;
Matches 124; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 2696 TGGATTGTGAACCCATCATCCGGAACAACTGTGGATAACTTATCGTTTCCGAATACAA 2755
Db 192 TAGAAGTGGCAACATACTCCCGGGCACTGTGTGGTGAAGATTGGCCATCCCAACCGA 251
QY 2756 ATTCGATTTTCTTGGCATCTCATCATGGTGTCTTGGTACATCTCGTCCAGGACATT 2815
Db 252 GTTTCATTTTACCTGTGCAGCCATGTGGCATTCCAGGGAACAAGCGTCTCGCCATT 311
QY 2816 CACTGTTATGATGAGTAAGGAATGAGCCAAAGATGAAGTCTATAAAATGACCTACCG 2875
Db 312 TCATGTTCTGTGGATGAGAACAAATTTACGGCTGATGGGTTGCAAACTCTCACCACAA 371
QY 2876 ACTTGCTTTCTCTCTGCTAGATGCGAAACCCATCTCGTTGCTGTTCCGTTTCATT 2935
Db 372 CTTGTTTACAGTATGCCAGGTGCACAGCTCAGTATCAATCGTTCTCTCTGCATCTA 431
QY 2936 TGCTCATTTATCA 2948
Db 432 TGCTCATCTGGCA 444

RESULT 15

US-11-218-305-22485
; Sequence 22485, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22485
; LENGTH: 965
; TYPE: DNA

Search completed: July 5, 2006, 23:20:24
Job time : 386 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:24:51 ; Search time 197 Seconds
(without alignments)
2367.313 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 MGSNPPELEKGFYRHSLDPE.....RHMEHFLQTNVYQMGSPA 1020

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*
10: Geneseq2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5349	100.0	1020	AAU01856	Aau01856 C. elegans
2	604.5	11.3	997	ADQ80699	Adq80699 Arabidops
3	577	10.8	857	ADJ94697	Adj94697 Human euk
4	577	10.8	857	ADJ94697	Adj94697 Human euk
5	577	10.8	857	AEC01709	Aec01709 Human EIP
6	577	10.8	857	AEF73789	Aef73789 Human Arg
7	575.5	10.8	969	ABO43145	Abo43145 A. thalia
8	575.5	10.8	969	ADC46703	Adc46703 Thalecres
9	575.5	10.8	969	ADI43621	Adi43621 Plant tra
10	575.5	10.8	969	ADO02339	Ado02339 Thalecres
11	566	10.6	859	ABP64718	Abp64718 Human pro
12	566	10.6	859	AEF73791	Aef73791 Human Arg
13	566	10.6	860	ABG97473	Abg97473 Human NAA
14	566	10.6	860	ADJ94698	Adj94698 Human euk
15	560.5	10.5	950	ABB67046	Abb67046 Drosophil
16	560.5	10.5	984	ABB62573	Abb62573 Drosophil
17	560.5	10.5	984	ABB67045	Abb67045 Drosophil
18	560.5	10.5	984	ABG70016	Abg70016 Larval vl
19	555.5	10.4	855	ADJ94700	Adj94700 Human euk
20	555.5	10.4	860	AAM80292	Aam80292 Human pro
21	555.5	10.4	860	AAB94291	Aab94291 Human pro
22	555.5	10.4	860	ABG72612	Abg72612 Human cvt
23	555.5	10.4	860	ADH77275	Adh77275 Human PAZ

24	555.5	10.4	860	10	AEF73790	Aef73790 Human Arg
25	555	10.4	861	10	AEF73792	Aef73792 Human Arg
26	555	10.4	924	7	ADJ69189	Adj69189 Human hea
27	555	10.4	924	8	ADJ94699	Adj94699 Human euk
28	554.5	10.4	860	6	ABU11762	Abu11762 Human MDD
29	550.5	10.3	858	9	ADM42496	Adm42496 Human euk
30	549.5	10.3	816	6	AAE37044	Aae37044 Human nuc
31	542.5	10.1	782	4	AAB93139	Aab93139 Human pro
32	529.5	9.9	585	8	ADP55696	Adp55696 Human PRO
33	523.5	9.8	990	8	ADO61721	Ado61721 Transcrip
34	523.5	9.8	990	8	ADQ80705	Adq80705 Arabidops
35	519	9.7	1048	3	AAG42118	Aag42118 Arabidops
36	519	9.7	1048	6	ADB17512	Adb17512 A. thalia
37	519	9.7	1048	8	ADP43623	Adp43623 Plant tra
38	519	9.7	1048	8	ADQ80695	Adq80695 Arabidops
39	519	9.7	1048	9	AEC75736	Aec75736 Arabidops
40	515	9.6	868	3	AAG42119	Aag42119 Arabidops
41	515	9.6	924	9	ADY65088	Ady65088 S. mansoni
42	510.5	9.5	1013	8	ADQ80701	Adq80701 Arabidops
43	509	9.5	1064	3	AAG52593	Aag52593 Arabidops
44	505	9.4	884	3	AAG52594	Aag52594 Arabidops
45	501	9.4	906	3	AAG51563	Aag51563 Arabidops

ALIGNMENTS

RESULT 1
AAU01856
ID AAU01856 standard; protein; 1020 AA.
XX
AC AAU01856;
XX
DT 07-SEP-2001 (first entry)
XX
DE C. elegans RNA interference pathway protein RDE-1.
XX
KW RNA interference; RNAi; RDE-1; genetic interference; antibody; dsRNA;
KW double-stranded RNA-dependent gene silencing.
XX
OS Caenorhabditis elegans.
XX
PN WO200129058-A1.
XX
PD 26-APR-2001.
XX
PF 13-OCT-2000; 2000WO-US028470.
XX
PR 15-OCT-1999; 99US-0159776P.
PR 30-MAR-2000; 2000US-0193218P.
XX
(UYMA-) UNIV MASSACHUSETTS.
(CARN-) CARNEGIE INST WASHINGTON.
XX
Mello CC, Fire A, Tabara H, Grishok A;
WPI; 2001-316239/33.
N-PSDB; AAS03282, AAS03283.
XX
Novel RNA interference pathway genes and their protein products involved in mediation of genetic interference, useful for modulating and studying regulation of RNA interference pathway.
XX
Claim 3; Fig 6; 76pp; English.

The sequence represents the RNA interference (RNAi) pathway protein RDE-1. RDE-1 and RDE-4 are involved in the pathway mediating double-stranded RNA-dependent gene silencing (genetic interference) RDS-1 and RDS-4 protein is useful for preparing an RNAi agent, by incubating a dsRNA in the presence of the proteins. The prepared RNAi agents can be used as sequence-specific interfering agents for targeted genetic interferences. The nucleic acids are useful for studying the regulation of RNAi pathway and to generate knockout strains of animals such as C.elegans. RDE-1 and

CC RDE-4 genes and their products are useful for modulating RNAi pathway
CC activity. The polypeptides are useful for generating and testing
CC antibodies specific for the polypeptides which are useful for studying
CC the RNAi pathway in C.elegans and other organisms. RNAi pathway genes are
CC useful for mediating specific processes, e.g. a gene that mediates dsRNA
CC uptake by the cells is useful for transporting other RNAs into cells or
CC for facilitating entry of agents such as drugs into cells. RNAi pathway
CC mutant strains (rde-1, rde-4 mutants) are useful in genetic screens to
CC identify additional RNAi pathway components
XX Sequence 1020 AA;

Query Match 100.0%; Score 5349; DB 4; Length 1020;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNPFPELEKGFYRHSLSLDPPEMKWLPARPTGKDGKFEYKVKVLLVNVWFSSKIYDREYVE 60
DB 1 MSSNPFPELEKGFYRHSLSLDPPEMKWLPARPTGKDGKFEYKVKVLLVNVWFSSKIYDREYVE 60

QY 61 YEVKMTKEVLRNPKPKPPKTEIPIPRAKLFWOHLRHEKKQTFILEDYVDFSKDTVY 120
DB 61 YEVKMTKEVLRNPKPKPPKTEIPIPRAKLFWOHLRHEKKQTFILEDYVDFSKDTVY 120

QY 121 SYCRLNTVTSKMLVSEKVKVSKDEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
DB 121 SYCRLNTVTSKMLVSEKVKVSKDEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180

QY 181 EANSYKFKLVNMTQKRYAPVFNBEIKVQFARNFYDNNISILRVPSFDPNRFQESLE 240
DB 181 EANSYKFKLVNMTQKRYAPVFNBEIKVQFARNFYDNNISILRVPSFDPNRFQESLE 240

QY 241 VAPRIEAFGIYIGIKELPDGEPVLFNFAIVDLKFNAPQMSLLDYLLIVDPQSCNDDVR 300
DB 241 VAPRIEAFGIYIGIKELPDGEPVLFNFAIVDLKFNAPQMSLLDYLLIVDPQSCNDDVR 300

QY 301 KDLTKLMAGKMTIQOAPRIRQLLENLKLKCAEVWNNEMSLRTERHLTFIIDLCEENSL 360
DB 301 KDLTKLMAGKMTIQOAPRIRQLLENLKLKCAEVWNNEMSLRTERHLTFIIDLCEENSL 360

QY 361 VYKVGKSDRGNAKKYDTTLFKIYEENKKFIEFPHLPLVKYKGAKEVAVPMEHLEVHE 420
DB 361 VYKVGKSDRGNAKKYDTTLFKIYEENKKFIEFPHLPLVKYKGAKEVAVPMEHLEVHE 420

QY 421 KPQRYKNRIDLVMQDKFLKRAIRKPHDYKENTLKMELDPFSEBELNPFVERGLCSKLOM 480
DB 421 KPQRYKNRIDLVMQDKFLKRAIRKPHDYKENTLKMELDPFSEBELNPFVERGLCSKLOM 480

QY 481 IECPGKVLKEPMLVNSVNEIQMTPIVIRGFQSKQLNVVPEKELCCAVFVNNETAGNPCLE 540
DB 481 IECPGKVLKEPMLVNSVNEIQMTPIVIRGFQSKQLNVVPEKELCCAVFVNNETAGNPCLE 540

QY 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
DB 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQIMYDATKNEYAFYKNCITLNTGIGRFEIA 600

QY 601 ATEAKNMFERLPDKEQKLMFIIISKRLQNALYGFVKYCHDHTIGVANQHITSETVTKALA 660
DB 601 ATEAKNMFERLPDKEQKLMFIIISKRLQNALYGFVKYCHDHTIGVANQHITSETVTKALA 660

QY 661 SURHEKSGKRIFYQALKNALKGINQELDWSAEISPEKERRKTMPLTMVYGIDVT 720
DB 661 SURHEKSGKRIFYQALKNALKGINQELDWSAEISPEKERRKTMPLTMVYGIDVT 720

QY 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQEECRPGERAVAHGRERTDILEAKFVKL 780
DB 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQEECRPGERAVAHGRERTDILEAKFVKL 780

QY 781 LREFAENNDRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSVKQFMSEDEGDEPEPKYT 840
DB 781 LREFAENNDRAPAHIVVYRDGVSDSEMLRVSHDELRSLSKSVKQFMSEDEGDEPEPKYT 840

QY 841 FIVIQKRNHTRLLRMEKDKPVVKNKDLTPAETDVAVAVKQWEEDMKESKETGI VNPSSG 900

DB 841 FIVIQKRNHTRLLRMEKDKPVVKNKDLTPAETDVAVAVKQWEEDMKESKETGI VNPSSG 900
QY 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMSSQDEYVYKMTYGLAFLSARC 960
DB 901 TTVDKLIIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMSSQDEYVYKMTYGLAFLSARC 960

QY 961 RKPISLPPVPHVYAHLSCEKAKELYRTEYKHYIGDYAOPRTRHEMHEFLQTNVKYPGMSFA 1020
DB 961 RKPISLPPVPHVYAHLSCEKAKELYRTEYKHYIGDYAOPRTRHEMHEFLQTNVKYPGMSFA 1020

RESULT 2
ADQ80699
ID ADQ80699 standard; protein; 997 AA.
XX
AC ADQ80699;
XX
DT 21-OCT-2004 (first entry)
XX
DE Arabidopsis thaliana TFL1-binding protein #7.
XX
KW TFL1-binding protein; plant growth control; biotechnology;
KW fishing industry; screening.
XX
OS Arabidopsis thaliana.
XX
PN JP2004208572-A.
XX
PD 29-JUL-2004.
XX
PF 27-DEC-2002; 2002JP-00381220.
XX
PR 27-DEC-2002; 2002JP-00381220.
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX
XX WPI; 2004-538191/52.
DR N-PSDB; ADQ80698.
XX
PT Novel AGO protein, comprising PA2 and Piwi domain, specifically binding
PT to TFL1 protein that is derived from Arabidopsis thaliana, useful for
PT controlling growth phase of plant.
XX
PS Disclosure; SEQ ID NO 14; 104pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC Arabidopsis thaliana proteins which specifically bind to TFL1 protein.
CC The amino acid and coding sequences of the invention are useful for
CC screening TFL1 family proteins. The DNA and protein sequences of the
CC invention are also useful for controlling the growth phase of a plant,
CC and as a reagent for research in biotechnology and fishing industry. The
CC present amino acid sequence represents an Arabidopsis thaliana TFL1-
CC binding protein of the invention.

Sequence 997 AA;
Query Match 11.3%; Score 604.5; DB 8; Length 997;
Best Local Similarity 23.8%; Pred. No. 1.8e-41;
Matches 242; Conservative 164; Mismatches 353; Indels 259; Gaps 41;

QY 38 KVVLLLVNVPFSSKIYDREYVEYKVKMKGKPKPPKTEIPIPRAKLPMOHL 97
DB 158 KVVVRAN--HFLVQVADRDLHYHVSINPEVSK----- 190

QY 98 RHEKKQTDILEDYVDFSKDTVYSYCRLNTVTSKMLVSEKVKVSKDE---KKDEKLEKK 154
DB 191 -----TVNRNVNMLKVN---YKDSHLGGKSPAYD-GRK 220

QY 155 ILYTM--ILTYRKKFHLNFSRENPEKDEANRSYKF-LKNV-----MTQKVRVA 200
DB 221 SLYTAGPLPFPSKEFVNVNLAEKRADGSGGKORPFKVAVKNVTSITDLYQLQQLDRKQREA 280


```

Db      311  KKYNNLQKYPHLPCLQVQEQKHLYPLEVNCI--VAGQRCIKKLTNDNOTSTWIKATARS 369
QY      445  PHDYKENTLKMKLKELDFSSSEELNVERFGLCSKLQMIQPCGKVLKEPMLVNSVNEQIKWT 504
Db      370  APDRQEEISRLMKNASYNLDP--YIQEFGIKVKDDMTVTGRVLPAPILQYGGNRRAIAT 427
QY      505  P-----VIRGPEKQL--NVVPEKELCCAVPVNNEAGNPLEENDDVVKFYTELI----- 552
Db      428  PNQGVMDMRG---KQFYNGIBIKWAIACFAPQK-----QCREE--VLKNFTDQLRKISKD 478
QY      553  GGCKFRGIRIGANENRGAQSIIMVDTAKNEYAFYKNCI--VAGQRCIKKLTNDNOTSTWIKATARS 612
Db      479  AGMPQGGQPCFKYAGGADSV-----EPMFRHLK 507
QY      613  DKEQKVLMPFIISKRQLNAYGFVKGHCHTIGVANOHTITSETVTKALASLRHEKSKRIF 672
Db      508  NTYSGQLIIVILPGKTPVYAEVKRGVDTLLGMATQCVQKNVVKT-----SPQTL 558
QY      673  YQALKINAKLGGINQELDWSIAEISPEEKERRKTMPLTMVIGDVTHTPTSYSGIDYSI 732
Db      559  SNLCLKINVKLGGINNIL-----VPHQSAVFQPPV--IFLGADVTHPPAGDGKGPSI 609
QY      733  AAVVASINPGGTYIRNMIVTQECRPGERAVAHGRERTDILE--AKFYK--LLREFAENNID 789
Db      610  TAVGSMDAHPSRYCATRVQ-----RPRQELIEDLSYMWRELLIQFYKSTR 656
QY      790  NPRAHIVVYRDGVSDSEMLRVSHDELRSKSEVKQFMSERDGEDPEPKYTFIVIQKRN 849
Db      657  FK--PTRIIPYRGVPEGLPQILHVELLAIRDACIKL-----EKDYQPGITVIVVQGRHH 710
QY      850  TLLRRMEKDKPVNKKLTPTAETDAVAANKWEEDEKESKETGIWNSSGTTVDKLIUS 909
Db      711  TRLF--CADKNERIGKSGNIPA-----GTFVDTNITH 740
QY      910  KYKDFPFLASHHGVLGTRPGHYTMVDDKGMQDEVKYQWYGLAFLSARCKRPISLPVP 969
Db      741  PEPDFYLCSHAGIOGTRSPSHYYVLDNDRRTADELQILTQLCHTYVTRCTRSVSIAP 800
QY      970  VHYAHLSCBKAKELRYTYKEHYGD 994
Db      801  AYYARLVAPRAR--YHLVDKEHDSGE 824

```

RESULT 4

ADJ94697
ID ADJ94697 standard; protein; 857 AA.

AC ADJ94697;

DT 06-MAY-2004 (first entry)

XX Human eukaryotic translation initiation factor 2C1 (eIF2C1) protein.

DE cytosolic; antiinflammatory; virucide; immunosuppressive; tumour;

KW inflammatory; infectious disease; viral infection; degenerative;

KW autoimmune; gene therapy; argonaute family;

KW eukaryotic translation initiation factor 2C1; eIF2C1; human.

OS Homo sapiens.

XX WO2004007718-A2.

XX 22-JAN-2004.

XX 10-JUL-2003; 2003WO-BP007516.

XX 10-JUL-2002; 2002EP-00015532.

XX 23-AUG-2002; 2002EP-00018906.

XX (PLAC) MAX PLANCK GBS FOERDERUNG WISSENSCHAFTEN.

XX Tuschl T, Martinez J, Patkaniowska A, Urlaub H, Luehrmann R;

XX WPI: 2004-122948/12.
DR N-PSDB; ADJ94703.

XX New single-stranded RNA molecule having a length from 14-50 nucleotides,
PT useful for preventing or treating tumor, inflammatory, infectious, e.g.
PT viral infections, degenerative and autoimmune diseases.

XX Example; Fig 14; 73pp; English.

XX The invention relates to a novel single-stranded RNA molecule having a
CC length from 14-50 nucleotides where at least 14-20 of the 5'-most
CC nucleotides are substantially complementary to a target transcript. The
CC RNA molecule of the invention demonstrates cytostatic, antiinflammatory,
CC virucide and immunosuppressive activities and may be useful for
CC inhibiting the expression of a target gene in vitro or in vivo,
CC preferably for preventing or treating diseases associated with the
CC overexpression of at least one target transcript. The diseases may be
CC selected from tumour diseases, inflammatory diseases, infectious diseases
CC such as viral infections, degenerative diseases and autoimmune diseases.
CC Furthermore, the molecules of the invention may be utilised during gene
CC therapy. The current sequence is that of the human eukaryotic translation
CC initiation factor 2C1 (eIF2C1) protein of the invention.

XX Sequence 857 AA;

Query Match 10.8%; Score 577; DB 8; Length 857;

Beat Local Similarity 23.0%; Pred. No. 3.1e-39;

Matches 227; Conservative 171; Mismatches 365; Indels 222; Gaps 37;

QY 38 KVVLLVNVWFSSKIYDREYVEYVKMTKEVLNKKPGKPKTEIPIPDRAKLFQWHL 97

Db 34 KPIKLLANFYFVIDPKIDVHYEVDIKPKD-----CPRVNRVE-----YMW 77

QY 98 RHEKQKTDIFLEDY--VFDEKDTVYVSVCLNTVTSMKVSEKV---VKDSKSKDEKJLE 152

Db 78 QHFKPQ---IFGDRKPVYDGKNIY-----TVTALPIGNRVDPEVTIPGEGKQ-RIFK 127

QY 153 KKILYTMILTTRKKPHLNFPSRENPEKDEANRSYKFLKNVMTQKRYAPFVNEEIKVQPA 212

Db 128 VSIKWLAIVSWRMLHEALVSGQIPVPLESVQALDVAMRHLS--MRYTPV----- 175

QY 213 KNFVYDNNISILRVPSFHDNRFQESLEVAPRIEAWFGIYIGIKELFQCEPVLNFAIVDK 272

Db 176-----GRSFPSPGEGYHP-----LGGREWFQGHQSVRPAM-WKMLNIDVSAT 220

QY 273 LFTYNAKMSLLDYLLIYDPOSCNDDVRKDLTKLMAGQMTTIRQAARPIROLLENLKK 332

Db 221 AFYKA--QPVEFCMEVLDIRNIDQPK-----PLTDSQVRVFTKEIKGLKVE 266

QY 333 ---CAEVDN-----EMSLRTERHLTFLDLCEENSLVYKVTGKSDRGNKKYDTTLFKIY 385

Db 267 VTHCGQMKRKYRVNCVTRRPASHQTFPLQLESGQTV-----ECTVAQYP 310

QY 386 EENKFF-IEFPHLPVKKVSGAKYAVPMEHLEVHEKQRYKNRIDLVNQDKFLKRAKTR 444

Db 311 KQKNLQKYPHLPCLQVQEQKHLYPLEVNCI--VAGQRCIKKLTNDNOTSTWIKATARS 369

QY 445 PHDYKENTLKMKLKELDFSSSEELNVERFGLCSKLQMIQPCGKVLKEPMLVNSVNEQIKWT 504

Db 370 APDRQEEISRLMKNASYNLDP--YIQEFGIKVKDDMTVTGRVLPAPILQYGGNRRAIAT 427

QY 505 P-----VIRGPEKQL--NVVPEKELCCAVPVNNEAGNPLEENDDVVKFYTELI----- 552

Db 428 PNQGVMDMRG---KQFYNGIBIKWAIACFAPQK-----QCREE--VLKNFTDQLRKISKD 478

QY 553 GGCKFRGIRIGANENRGAQSIIMVDTAKNEYAFYKNCI--VAGQRCIKKLTNDNOTSTWIKATARS 612

Db 479 AGMPQGGQPCFKYAGGADSV-----EPMFRHLK 507

QY 613 DKEQKVLMPFIISKRQLNAYGFVKGHCHTIGVANOHTITSETVTKALASLRHEKSKRIF 672

Db 508 NTYSGQLIIVILPGKTPVYAEVKRGVDTLLGMATQCVQKNVVKT-----SPQTL 558

Qy 970 VHYAHLSCAKELYRTRYKEHYIGD 994
: || | : | : || | :
Db 801 AXYARLVAFRAR-YHLVDKEHDSGE 824

RESULT 6

AEF73789
ID AEF73789 standard; protein; 857 AA.

AC AEF73789:

DT 06-APR-2006 (first entry)

Human Argonaute protein hAgo1.

KW Argonaute; hAgo1; RNA interference; RNAi; drug screening;
 JW X-ray crystallography; gene silencing; protein structure;

OS Homo sapiens.

Key	Location/Qualifiers
FH Domain	230. .350
FT	/label = PAZ_domain
FT	
FT Domain	581. .830
FT	/label = PIWI_domain
FT	

PN WO2006015258-A2.

PD 09-FEB-2006.

28-JUL-2005: 2005WO-US027084.

28-JUL-2004: 2004US-0592297P.

117 XX

PI Joshua-Tor L, Song J, Hannon GJ, Liu J, Carmell MA, Rivas F;
PI Marsden C;
XX
DR WPI: 2006-155768/16.

Crystalline Argonaute useful for identifying agent that binds Argonaute protein, utilized for treating pancreatic cancer, treating hepatitis C infection, cancer or inflammatory diseases.

PS Disclosure; SEQ ID NO 1; 215pp; English.

The present sequence is that of human Argonaute protein hAgo1. Argonaute proteins are involved in RNA interference (RNAi). The invention provides a crystalline Argonaute protein, which comprises (i) an N-terminus, middle and PIWI domain which form a crescent-shaped base, and (ii) a PAZ domain, which is positioned above the crescent-shaped base, resulting in a cleft between the crescent-shaped base and the PAZ domain. The structure of the full-length Argonaute protein AGF73751 from the archaeobacterium *Pyrococcus furiosus* (PfiAgo), as determined by x-ray crystallography to 2.25 Angstrom resolution, is provided. The invention also provides an isolated complex comprising an Argonaute protein and a single-stranded RNA hybridized to its target nucleic acid. In certain embodiments, the single-stranded RNA is bound to the PAZ domain of the Argonaute protein, and may further interact with the crescent-shaped base of the Argonaute protein. Also claimed are: a method of determining the three-dimensional structure of an Argonaute protein or its mutant, derivative, variant, analog, homolog, sub-domain or fragment by alignment with the PfiAgo amino acid sequence to match homologous regions; a method of identifying an agent that binds an Argonaute protein by applying a three-dimensional molecular modeling algorithm to the atomic coordinates of an Argonaute protein to determine the spatial coordinates of the binding pocket of the Argonaute protein, and electronically screening the stored spatial coordinates of a set of candidate agents against the spatial coordinates of the Argonaute protein binding pocket; a computer-based method for the analysis of the interaction of a molecular structure


```
Db 416 IRLRFLTEKSEKTVQVFAKKNYVKQALPAIQTGSDTRPVYLPMLCQIDE-GQR 474
Qy 425 YKNRIDVMDQFKLRATRPHDYKENTLKMJELDFSSBELNPFVFGCLSKLQITECP 484
Db 475 YTKRLNEKQVTKATCQRPDP-RENSIKNLVVKNNYNDLS--KEFGMSVTTQLASIE 531
Qy 485 GKVLKEPMLVNSNEQIKMTFVIRGFBQKQLNVVPEKELCCAVFVNNETAGNCPLENDV 544
Db 532 ARVLPPLMKYHDSGKEMVNLRLG-----QWNMIDKK----- 564
Qy 545 VKFYTELIGCKFRGIRIGANENRGAOSIMYDANKNEYAFYKNCNTLMTGIRPEIAATEA 604
Db 565 -----MVNGAK---VTSWCEPKPQPAIPFISCPPEH-----TEALLD- 600
Qy 605 KNNFERLPDKEQKVMFIIISKQLNAYFGVGHYCDHTIGVANQHITSETVTTKALASLRH 664
Db 601 --IHKRAPGLQ-----LLVILPDPVTSYGKIKRICETELGIVSQCCQPRQVNLK----- 648
Qy 665 EKGSKRIFYQIALKINAKLGINQELDWSIAEISPEEKERRKTMPL-----TMYGVIDV 719
Db 649 ---NKQTMENVALKINKVTCGRNTVLN-----DAIRRNIPLTDRPTIIMGADV 694
Qy 720 THPTSYSIGIDYSIAAVASIN-PGQTIYRNMTVQ---EE-----CRPGERAVAHGR 767
Db 695 THPOGEDSFSIAAVASMDWFEINKIRGLVSAQAHEEIIQDLYKLVQDPQORGLVH-- 752
Qy 768 ERTDILEAKFVKLLREPAENNDNRAPAHIVVYVRDGVSDSEMLRVSHDELSLSEVKQFM 827
Db 753 --SGLIREHFAFRA-----TGQIPQRIIFYRDGVSEGOFSQVLLHEMTAIRKACNSLQ 805
Qy 828 SERGEDPEPKYFIVIOKRNHLLRMKDKPVPVNNKOLTPAETDVAAVAAVQWBEEDMK 887
Db 806 -----ENVPRVTVFVIOKRNHLLRFPPEHG-----NRDMT----- 836
Qy 888 ESKETGVNPSSTGVTDKLVSKVPDFFLASHHGLVGTSPGHYTVMYDDKMGSDDEVY 947
Db 837 --DKSGNIQP--GTVDVTKICHNEFDYILNSHAGIQGTSRPAHYHVLLENGFTADQLQ 892
Qy 948 KMTYGLAFLSARCKPISLPVPHVYAHLSCEKAKELYRTYKHEYIGDYAQPRTR 1001
Db 893 MLTNLCYTVARCTKSVSIVPPAYAHAAAFRA---RYTMESEMSDGGSSRSR 942

RESULT 9
AD143621
ID AD143621 standard; protein; 969 AA.
AC AD143621;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant transcription factor #683.
XX
KW transgenic; plant; enhanced tolerance to abiotic stress;
KW phosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; gene; ds.
XX
OS Arabidopsis thaliana.
XX
XX US2004019927-A1.
XX
XX 29-JAN-2004.
XX
XX 25-FEB-2003; 2003US-00374780.
XX
XX 18-APR-2001; 2001US-00837944.
PR
```

```
XX
PA (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (FINE/) PINEDA O.
PA (YUGG/) YU G.
XX
PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V,
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE,
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX
DR WPI; 2004-132245/13.
XX
XX New transgenic plant comprising a recombinant polynucleotide of any one
XX of more than 500 nucleotide sequences, useful in bioinformatic search
XX methods.
XX
PS Claim 1; SEQ ID NO 2084; 435pp; English.
XX
XX The invention describes a transgenic plant comprising a recombinant
XX polynucleotide of any one of more than 500 nucleotide sequences fully
XX defined in the specification or its complement. The method of the
XX invention can be used to produce a plant having altered traits such as:
XX enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
XX sensitivity; disease resistance; sugar sensing; early or late flowering;
XX altered flower structure, change in stem bifurcations, altered branching
XX pattern, reduced apical dominance, reduced trichome density; lack of
XX trichomes; increased ectopic trichome development; altered trichome
XX development; increased root hairs; altered stem morphology;
XX increased cell proliferation or cell differentiation; rapid development;
XX premature senescence; increased necrosis; increase in seedling or plant
XX size; decreased plant size; leaf morphology; seed morphology; seed
XX biochemistry; increase in root anthocyanins; increase in plant
XX anthocyanins, or alteration in light response or shade avoidance. The
XX transgenic plant, polynucleotides and polypeptides are useful in
XX bioinformatic search methods. This is the amino acid sequence of a plant
XX transcription factor, and an orthologue of Arabidopsis thaliana
XX transcription factors isolated in the invention, that can be used in the
XX creation of a transgenic plant with altered traits.
XX
SQ Sequence 969 AA;
```

```
Query Match 10.8%; Score 575.5; DB 8; Length 969;
Best Local Similarity 23.2%; Pred. No. 5e-39;
Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;
Qy 38 KKVLNVNFKFSKITYEYEVKMTKVLNKRKPGKPPKKTETIPIDRAKLFQHL 97
Db 158 KKVMVPAH-HFLVQVADRDLHYDVSINPEVISK----- 190
Qy 98 RHEKKTQDILEDYVDFDKTVVSVCLNTVTSKMLVSEKVKVQDS---KKDEKDLKK 154
Db 191 -----TVNRNVNKLIVKN---YKDSHLGGKSPAYD-GRK 220
Qy 155 ILTYM--ILTYRKKFHLNFSRENPEKDEANRSYKF-LKNV-----MTQKRYA 200
Db 221 SLYTAGLPFDSEKFEVFNLAERADGSSGKORPFKAVKNNVTSTLDYQLQOQFLDRKQEA 280
Qy 201 PFVNEIKVQFAKNFVYDNNNSILRVPSFHDNRFQSL-----EVAPRIEAWFGI 251
Db 281 PY--DTIQVLDVLRDKPNDYVSGRSP-----FHTSLGKDARDGELGDIYWRGY 333
```


CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phylloclax, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed morphology, increased root
 CC proliferation/cell differentiation, premature senescence, altered cell
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenly lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenly lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence represents a
 CC thalaeass transcription factor of the invention.
 CC
 XX Sequence 969 AA;

Query Match 10.8%; Score 575.5; DB 8; Length 969;
 Best Local Similarity 23.2%; Pred. No. 5e-39;
 Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;
 QY 38 KVVLLVNNFKPSKIDREYVEYVMTKEVLRNKGKPPKTEIPIDRAKLFQWHL 97
 DB 158 KVMVRAN--HFLVQVADRDLHYDVSVINPEVSK----- 190
 QY 98 RHEKKQTDILEDYVDEKDTVYVSVCLNATVTSKMLVSEKVKVCKDSE---KDEKLEKK 154
 DB 191 -----TVNRNVMKLLVKN---YKDSHLGSKSPAYD-GRK 220
 QY 155 ILVTM--ILTYRKPHLNFSPENPEKDEANRSYKE-LKNV-----MTQKRVRA 200
 DB 221 SLYTAGLPDSKEFVNLAERADSGGKORPFKAVKXVNTDLYQOQFLDRKQREA 280
 QY 201 PFVNEIKVQFAKVFYDNNLSILRVSPESHDPNRFQSL-----EVAPRIEAWFGI 251
 DB 281 FY--DTIQVLDVVLDRKPSNDYVSVGRSP-----FHTSLGKARDGKRGELGDIEYWRGY 333
 QY 252 YIGIKELFDGEFVFNPAIVDKLPFNAPKMSLLDYLIVDPQSCNDVVRKOLTKLWAGK 311
 DB 334 FQSLR--LTQGLUSLNDVRSARYP-----EPVVVTFDFISKLNIRDL--N 375
 QY 312 MTIRQAARPRIRQLLENLKLKCAEVMDNEMSLRHLTFDLCEENSLVYKVTGKSD-- 369
 DB 376 RPLRDSRLKVKVLTLLKVLH--WNGTSA-----KISGSSLP 415
 QY 370 ----RGNAAKVDITLLFKIYEENKPF-IEFPHLPLVVKVSGAKAYAVPMHELVHEKPKR 424
 DB 416 IRELRFLEDKSEKTVVQVFAEKYKRVKYQALPAIQGTGDTRPVYLPMLCQIDE-GQR 474
 QY 425 YKNRIDLVMQDFLKBATRPKDPHYKENTLKMELDFSSSEELNPFVERFGLCSKLQWIECP 484

DB 475 YTRLNEKQVYALLKATCQPPD--RENSIKNLVVKNNYNDLDS--KEFGMSVTTQLASIE 531
 QY 485 GKVLKEPMLVNSVNEQIKMTPTVIRGPQEKOLNVVPEKELCCAVFVNVNAGNCPLEENDV 544
 DB 532 ARVLPPLMKYHSGKEMVNPRLG-----QWNMDKK----- 564
 QY 545 VKFVTELLIGCKFRGIRIGANENRGAQSIYDATKNFYAFYKNCNTLNTGIGRFFIAATEA 604
 DB 565 -----MVNGAK---VTSWTCFEPKQPAIPFISCPPEH-----IEEALLD- 600
 QY 605 KMFPERLPDEKQVLMFIIISKQLNAYGFVKYCHDTHTIGVANOHITSETVTTKALASLRH 664
 DB 601 --IHKRAPGLQ-----LLVILPDVTGSYKIKRICEITELGIVSQCCQPRQVKNL----- 648
 QY 665 EKSGKRFYQIALKINAKLGINOELDWSBIAEISPEKERRKTMPL-----TWVVGTDV 719
 DB 649 ---NKQYMNVALKINKVTKGRVTNL-----DAIRRNIPLITDRPTIIMAGDV 694
 QY 720 THPTSYSGIDYSIAAVVASIN-FPGTIYRNMIVTQ-----EE-----CRGERAVAHGR 767
 DB 695 THPQGEDSSPSIAAVVASMDWPEINKYRGLVSAQAHRBEEIQDLYKLVDPQRLVH-- 752
 QY 768 ERTDILEAKPKVLLRFAENNDRAPAHIVYVYRDGSDSEMLRVSHDELKSLKSEVQFQM 827
 DB 753 --SGLIREHFIAFRRA-----TCQIPQRIIFYRDGVSQGVFQVLLHEMTAIRKACNSLQ 805
 QY 828 SERDGDPEPKYFVIVQKHNTRLRLRMKDKPVVKNKOLTPAETDVAAVAKVQWEDMK 887
 DB 806 -----ENVYPRVTFVIVQKHNTRLRPEQHG-----NRDWT----- 836
 QY 888 ESKETGINVPSSGTTVDKLVSKYKDFFLASHHGVLTGSRPGHYTVYDDKGMODEVY 947
 DB 837 --DKSGNIQD--GTVVDTKICHNEFDFYLSHAGIQGTSRPAHYHVLLENGPTADQLQ 892
 QY 948 KMTYGLAFLSARCEKPSLPPVHYAHLSCAKELVYRTYKHYIGDYAQPRTR 1001
 DB 893 MLTNLCYTYARCTKSIVPPAYAYALAFRA-----RYTMESEMSDGGSSRSR 942

RESULT 11
 ABP64718
 ID ABP64718 standard; protein; 859 AA.
 XX
 AC ABP64718;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human protein SEQ ID 378.
 XX
 KW Human; expressed sequence tag; EST; haematopoietic disorder;
 KW central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 XX
 OS Homo sapiens.
 XX
 XX WO200259260-A2.
 XX
 PD 01-AUG-2002.
 XX
 XX 16-NOV-2001; 2001WO-US042950.
 XX
 PR 17-NOV-2000; 2000US-00714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH;
 PI Ren F, xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI: 2002-590824/63.
DR N-PSDB; ABQ99304.
XX New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX
XX Claim 20; SEQ ID NO 378; 394pp; English.
XX
XX The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotential or pluripotential state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIFO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 859 AA;

Query Match 10.6%; Score 566; DB 5; Length 859;
Best Local Similarity 23.1%; Pred. No. 2.7e-38;
Matches 234; Conservative 160; Mismatches 37; Indels 250; Gaps 39;
QY 19 PEMKWLARTGKCDGKFEYKVLVLLNWPKESSKIYDREYVEYVQMTKEVLNRKPKPF 78
DB 17 PIQGYAFKPPRPDPFTSGRTIKLQANFFEMDIPKIDIVHYELDIKPEK-----C 66
QY 79 PKKTEIPIPRAKLFWQHLRHEKKQTDFILEDY--VFDEKDTVYSVCLNTVTTSKMLVSE 136
DB 67 PRVNRREIYEHM-----VQHFKTQ---IPGDRKPVFDGRKNLYTAMPLPIGRDKV---E 114
QY 137 KVVVKDSEKKDEKLEKKILYTMILYTKKFKHLNFSRENPEKDEANRSYKFLKNVMTQ- 195
DB 115 LEVTLPGEGKDR--IPKSIKWVSCVSLQALHDALSGRLFSVPFE---TIOALDVVMRHL 169
QY 196 -KVRVAPFVNEEIKVQFKNFVYDNNISILRVPESHDNRFEQSLVAPRIEAWFGIYIG 254
DB 170 PSMRYTPV-----GRSFTASEGCSNP-----LGGREVFWEFGTHS 205
QY 255 IKELFDGEPVLPFAIVDKLFPYNAKMSLDYLLIIVDPOSCNDDVRKDKLTKLMAGKMTI 314
DB 206 VRPSL-WKMWLNIDVSATFYKA--QPVIEFCEVLDPKSIIEQ-----QKPL 250
QY 315 ROAARPRIQLLENLKLK---CAEVWDN-----EMSLRTERHLTFLDCENSLVYKVTGK 367
DB 251 TDSQVRVKFTKEIKGLKVEITHCCQMKRKVRVCNVTRRPASHQTF-PLQESGQTVECT-- 307
QY 368 SDRGRNAKYDITLTKIYEENKFIIEFPHLPVKKVSGAKYAVPMVHEHLEVKPQRYKN 427
DB 308 -----VAQY-----FK--DRHKLVLRYPHLPCLQVQEQBKHTYLPLEVCNIVAGRCICK 355
QY 428 RIDLVNQDKFLKRAAT-RKPHDYKENTLQMKLELDFSSBELNFVERFGLCSKLQMBPCPK 486

DB 356 LTD--NOTSTMIRATARGAPDRQBEISIKLMSASFNTPD--YVREFGIMVXDEMTDVTGR 411
QY 487 VLKEPMLVNSVNEQIKMTPV--IRGFQEKQLNV-----VPEKELCCAPVNVNE 532
DB 412 VLQPPSILYGRNKAIATPVQGVWDMRNKQFHTGIEIKVWAIACAPAQKQ-CTEVHLKSP 470
QY 533 T-----AGNPCLEENDVWVKFYTELIGGCKFRGIRIGANENRGAQSI--MYD 576
DB 471 TEQLRKISRDAGMPIQGOQPCF-----CKYA-----QGADSVPEMPFR 506
QY 577 ATKNEYAFYKNCNTLMTGIRFEIAATEAKNMFERLPDKBEQKVLMIISIKRQLNAYGFVK 636
DB 507 HLKNTYA-----GLQLVVVILPGKTPVYAEVK 533
QY 637 HYCDHTIGVANQHTSETVTTKALSLRHEKSGKRIFYQIALAKINAKLGGINQELDWSEIA 696
DB 534 RVGDTVLGMATQCQVMKNV-----ORTPTQLSNLCLKINVKLGGVNNIL----- 578
QY 697 EISPEKERRKTMPLTMVYGVIDVTHPTSVSGIDYIAAVVASINPGTYYRNMTQBERC 756
DB 579 --LPQGRPPVFOQPV-IFLGADVTHPPAGDGKKPSIAAVVGSMDAHPNRYCATVRVQ-- 633
QY 757 RPERAVAHGERTDILEAKFVKLREFAENNDRNAPAHIVVYRDCVDSSEMLRVSHDEL 816
DB 634 -----HRQEIIQDLAAMVRELLIQFYVKSTRFK-PTRIIFYRDCVSGEQFOQLLHHEL 684
QY 817 RSLKSEVKQFMSERDGEPEPKYTFIVIQKHNTLRLRMEKDKPVWVKDLTPAETDVAV 876
DB 685 LAIREACIKL-----EKDYQPGITFIVQKRHHTLRF-CTDKNERVGSKNIPA----- 732
QY 877 AAVKQWEEEDMKESKETGIVNPSGGTVDKLIYVKYKFDFFFLASHRHGLVLTSPGHYVMY 936
DB 733 -----GTTVDTKITHPTFEFDYLCSHAGIQGTSRPSHHVLM 769
QY 937 DDKMGSDQEVYKMTYGLAFLSARCKKPISLPVVHYAHLSCBKAKELYNTYKEH 990
DB 770 DDNRFSDELQILTYQLCHTVYVTRSVSIPAPAYYAHVAPRAR-YHLVDKEH 822
RESULT 12
AEF73791
ID AEF73791 standard; protein; 859 AA.
XX AEF73791;
XX AC AEF73791;
XX DT 06-APR-2006 (first entry)
XX Human Argonaute protein hAgo2.
DE DE
XX X-ray crystallography; RNAi; drug screening;
KW X-ray crystallography; gene silencing; protein structure.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 232..352
FT Domain /label = PAZ_domain
FT Domain 583..832
FT Domain /label = PIWI_domain
XX WO2006015258-A2.
XX PN
XX PD 09-FEB-2006.
XX 28-JUL-2005; 2005WO-US027084.
PF 28-JUL-2004; 2004US-0592297P.
PR 29-JUL-2004; 2004US-0592269P.
XX (COLD-) COLD SPRING HARBOR LAB.
XX Joshua-Tor L, Song J, Hannon GJ, Liu J, Carmell MA, Rivas P;
PI Marsden C;

XX WPI; 2006-155768/16.

XX Crystalline Argonaute useful for identifying agent that binds Argonaute

XX protein, utilized for treating pancreatic cancer, treating hepatitis C

PT infection, cancer or inflammatory diseases.

PT

XX

XX Disclosure; SEQ ID NO 3; 215pp; English.

XX

CC The present sequence is that of human Argonaute protein hArgo2. Argonaute

CC proteins are involved in RNA interference (RNAi). The invention provides

CC a crystalline Argonaute protein, which comprises (i) an N-terminus,

CC middle and P1WI domain which form a crescent-shaped base, and (ii) a PAZ

CC domain, which is positioned above the crescent-shaped base, resulting in

CC a cleft between the crescent-shaped base and the PAZ domain. The

CC structure of the full-length Argonaute protein ABF73751 from the

CC archaeobacterium *Pyrococcus furiosus* (PfuAgo), as determined by x-ray

CC crystallography to 2.25 Angstrom resolution, is provided. The invention

CC also provides an isolated complex comprising an Argonaute protein and a

CC single-stranded RNA hybridized to its target nucleic acid. In certain

CC embodiments, the single-stranded RNA is bound to the PAZ domain of the

CC Argonaute protein, and may further interact with the crescent-shaped base

CC of the Argonaute protein. Also claimed are: a method of determining the

CC three-dimensional structure of an Argonaute protein or its mutant,

CC derivative, variant, analog, homolog, sub-domain or fragment by alignment

CC with the PfuAgo amino acid sequence to match homologous regions; a method

CC of identifying an agent that binds an Argonaute protein by applying a

CC three-dimensional molecular modeling algorithm to the atomic coordinates

CC of an Argonaute protein to determine the spatial coordinates of the

CC binding pocket of the Argonaute protein, and electronically screening the

CC scored spatial coordinates of a set of candidate agents against the

CC spatial coordinates of the Argonaute protein binding pocket; a computer-

CC based method for the analysis of the interaction of a molecular structure

CC with an Argonaute protein; a data array comprising the atomic coordinates

CC of an Argonaute protein; a computer-readable storage medium encoded with

CC the atomic coordinates of an Argonaute protein; an electronic

CC representation of a domain or binding site of the Argonaute protein; a

CC method for obtaining a crystal of an Argonaute protein; a method for

CC identifying an agent that modulates the activity of an RNAi construct by

CC identifying an agent that modulates the expression and/or activity of an

CC Argonaute protein; a method for identifying an agent that potentiates the

CC activity of an RNAi construct by identifying an agent that increases the

CC expression and/or activity of an Argonaute protein; a method of

CC identifying an agent that modulates the activity of an RNAi construct by

CC examining a change in Argonaute protein activity in the presence of a

CC candidate agent; a composition for targeted gene inhibition comprising an

CC agent that modulates the RNase activity of an Argonaute protein; a cell

CC line that overexpresses an Argonaute protein; an assay for identifying

CC nucleic acid sequences for conferring a particular phenotype in a cell; a

CC nucleic acid composition composed of a nucleic acid comprising an RNAi

CC construct and a nucleic acid encoding an Argonaute protein, where the

CC RNAi construct comprises a nucleic sequence encoding a single-strand

CC short interfering RNA (siRNA); a pharmaceutical composition comprising

CC the nucleic acid composition; and a cell expressing the nucleic acid

CC composition. The methods and compositions of the invention are useful for

CC enhancing the effectiveness of an RNAi therapeutic.

XX

XX Sequence 859 AA;

SQ

Query Match 10.6%; Score 566; DB 10; Length 859;

Best Local Similarity 23.1%; Pred. No. 2.7e-38;

Matches 234; Conservative 160; Mismatches 370; Indels 250; Gaps 39;

QY 19 PEMKWLARPTGKCDGKFKYKVLVNVFKFSSKIYDREYEVYEVMTKEVLNRKPKGP 78

DB 17 PIQGYAFKPPDPDFTSGRTIKLQANFFEMDIPKIDIVHYELDIRPEK-----C 66

QY 79 PKKTEIPIDRAKLFWQHLRHEKKQTDFILEDY--VFDEKQTVYSVCLRNVTWTSKMLVSE 136

DB 67 PRVNRNREIHEM-----VQHFKTQ---IFGDKPVPFDGKRNLYTAMPLPIGRDKV---E 114

QY 137 KVVYKQDEKDEKLEKKILYTMILYRKFKHLNFSRENPEKDEANRSYKFLKNVMTQ- 195

Db 115 LEVTLFEGEKDR--IFKVISIKWVSCVSLQALHDALSGRLPSVPFE-----TIOALDVNRHL 169

QY 196 -KRYAPFVNEEIKVOPAKNFVYDNNISILRVPSFSDHPNRFPEQSLEVAPRIEAMFGIYIG 254

Db 170 PSMRYTPV-----GRSFFTASBGCSNP-----LGGGREVWFGFHQS 205

QY 255 IKELFDGEPVNFVVDKLPYNAPKMSLLLDYLLVDPQSCNDDVDRKDLTKTKLMAGKWTI 314

Db 206 VRPSL-WKMLNLDVSATAFYKA--QPVIEFVCEVLDFKSI EQ-----OKPL 250

QY 315 ROAARPRIOLLENLKLK---CABVDN---BMSRLTERHLTPFLDICEENSLVVKYTKG 367

Db 251 TDSQRVKFTEIKGLKVEITHCGMKKRYVCNVRPASHQTP-PLQEGSGQVECT-- 307

QY 368 SDRGRNAKYDITLTKIYEENKKFIEPPHPLPVKVKSGAKAYAVPMHEHLEVHEKPYKYN 427

Db 308 -----VAQY-----FK--DRHLVLRYPHLPCLVQGOEQKHTYLPVLEVCNVAGORCIKK 355

QY 428 RIDLVMDQKFLKAT-RKPHDYKENTLKMELKDFPSSEELNFVERFGLCSKLQMIKPGK 486

Db 356 LTD--NQTSMIRATARSAPDQSEISKLMRSASFNTDP--YVREFGIMVKDEMTDVTGR 411

QY 487 VLKEPMLVNSVQIKMTVPV--IRGFOEKOLNV-----VPEKELCCAFVFNVE 532

Db 412 VLQPSILYGGRNKAIATPVQGVMDNRNKQFHTGIEIKVWAIACFAPQRO-CTEVHLKSP 470

QY 533 T-----AGNPLEENDVVVKFYTELIGCKPGRGIRIGANENRGAOSI--MYD 576

Db 471 TEQLRKISDAGMPIQGPCF-----CKYA-----QGADSVEPMFR 506

QY 577 ATKNEYAFYKNCITLNTGIRGFEIATAEKNMFERLPDKOEKVLMPFIISIKRQLNAYGFVK 636

Db 507 HLKNTYA-----GLQLVVVILPKTPVVAEVK 533

QY 637 HYCDHTIGVANOHTSETVTKALASLRHKGSKRIFVQIAKINAKLGGINOELDWSEIA 696

Db 534 RVGDTVLGATQCVQMKNV-----ORTPTLSNLCKLNKVLGGVNNIL----- 578

QY 697 EISPEEKERRKMTPLTMYGIDVTHTPSYSGIDYSIAAVVASINPGGTIYRNMIVTQEEC 756

Db 579 --LPQGRPPVFQGPV-IFLGADVTHPPAGDGKKPSIAAVVGSMDAHPNRYCATVRVQQ-- 633

QY 757 RGERAVAGRENTDLEAKFVKLLREPAENNDNRAPAIIVTRDGVSDSEMLRVSHDEL 816

Db 634 -----HREIIOQLAAVRELLIOFYKSTRFK-PTRIIFYRDGVSEGGQQOVLHLEL 684

QY 817 RSLKSVKQFMSERDEGPEPKYTFIVIOKRHNTRLRMEKDKPVVNKDLTAEITDVAV 876

Db 685 LAIREACIKL-----EKDQPGITFIVOKRRHRLFP-CTDKNRVKGSGNIPF----- 732

QY 877 AAVKQWEEDKESKETGIYVNPSSGTTVDKLIIVSKYKFPDFFLASHHGVLGTSRFGHYTVMY 936

Db 733 -----GTTVDTKITHPTTEFDYFLCSHAGIQGTSRFSHYHVLW 769

QY 937 DKGMSQDEYKMTYGLAFLSARCCKPISLPVPVHYAHLSCSEKAKELYRTYKEH 990

Db 770 DDNRFSSEDLQILTYQLCHTYVVRCTRSVSIPAPAYAHLVAFRAR-YHLVDKEH 822

RESULT 13

ABG97473

ID ABG97473 standard; protein; 860 AA.

XX

AC ABG97473;

XX

DT 16-DEC-2002 (first entry)

XX

DE Human NAAP7 from INCYTE no.1725129CD1.

XX

KW Human; nucleic acid associated protein; NAAP; cancer;

KW cell proliferative disease; cancer; atherosclerosis; hepatitis;

KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;

KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;

KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.
OS Homo sapiens.
XX WO200272630-A2.
XX 19-SEP-2002.
XX 07-FEB-2002; 2002WO-US003844.
XX 09-FEB-2001; 2001US-0268118P.
XX 21-FEB-2001; 2001US-0270963P.
XX 22-FEB-2001; 2001US-0270858P.
XX 23-FEB-2001; 2001US-0271194P.
XX 07-MAR-2001; 2001US-0274071P.
XX 12-APR-2001; 2001US-0283496P.
XX 09-NOV-2001; 2001US-0344650P.
XX (INCY-) INCYTE GENOMICS INC.
XX Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Wallia NK, Thangavelu K;
PI Elliott VS, Marquis JP;
XX WPI; 2002-723320/78.
DR N-PSDB; ABS78717.
XX
XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX
XX Claim 1; Page 144-146; 162pp; English.
XX
XX The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
CC screening for a compound that is effective as an ant/agonist or modulator
CC of NAAP, generating an expression profile of a sample containing the
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridizable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
CC other diseases and disorders listed in the specification. These are also
CC useful in assessing the effects of exogenous compounds on the expression
CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The microarray is useful in monitoring
CC or measuring protein-protein interactions, drug-target interactions, and
CC gene expression profiles. The present sequence represents an NAAP protein
XX
XX Sequence 860 AA;

Query Match 10.6%; Score 566; DB 5; Length 860;

RESULT 14

Best Local Similarity 23.1%; Pred No. 2.7e-38;
Matches 234; Conservative 160; Mismatches 370; Indels 250; Gaps 39;
QY 19 PEMKWLARPTGKDGKFEYKVKLLLVNMFKFSKSKYDREYVEYVMTKEVLNRKPGKPF 78
DB PIQGVAFKPPRPDPFGTSGRTIKLQANFEMDIPKIDIVHYELDIKPEK-----C 67
QY 79 PKTEIPIDRAKLFQWHLRHEKKQTDFILEDY--VFDEKDTVYSVCLRTNTVTSKMLVSE 136
DB PRVNVREIVEHM-----VQHEKTO---IFGDRKPVFDGRKNLYTAMPLPIGRDKV---E 115
QY 137 KVVKKDSEKKDSEKLEKILYTMILTYRKPHLNSRENPEKDEBENRYSYKFLKNVQ-- 195
DB LEVTLPGEGKDR--IFKVISIKWVSCVSLQALHDALSGRLPSVPFE--TIQALDVVNRHL 170
QY 196 -KVRYAPFVNEEIKVQAKNFYVDNNSILRVSPESHDPNRFQSLLEVAPRIEAWFGIYIG 254
DB PSMRYTPV-----GRSPFTASEGCSNP-----LGGGREVWFQFHQS 206
QY 255 IKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLLVDPQSCNDDVVRKDLTKLMAGKMTI 314
DB VRPSL-WKQMLNIDVSATAFYKA--QPVIEFVCEVLDFKSBEQ-----QKPL 251
QY 315 ROAARPRIOLLENLKLK---CAEYWDN---EMSLTERHLTLFLDLCEBSLVYKVTGK 367
DB TDSQVRKTKKEIKGLKVEITHCGQMKRKYVVCNVTTRPASHQTF-PLQESGQTVECT-- 308
QY 368 SDRGRNAKKYDITLFLKIYEENKCFIEFPHLPVLVKYSGAKYAVPMHEHLEVHEKPPORYN 427
DB -----VAQY----FK--DRHKLVLRYPHLPCLQVQEQKHTYLPLEVCNIVAGQCIKK 356
QY 428 RIDLVQDKFLKRAI-RKPHDYKENTLXMKLELDFSSBELNFEVERGLCSKLMQISCPGK 486
DB LTD--NOTSTWIRATARSAPDRQESIKLMRSASFNTDP--YVREFGINVMKDEMTDVTGR 412
QY 487 VLKEPMLVSNVNEQIKMTPV--IRGFOBKQLNV-----VPEKELCAVAVVNE 532
DB VLQPSILYGGNKAIATFVQGVDMRNKQFHTGIEIKVWATACAPAPQ--CTEVLKSF 471
QY 533 T-----AGNCPLEENDVVKFYTELIGGCKFRGIRIGANENRGAQSI---MYD 576
DB TEQLRKISRDAQMPIQGQPCF-----CKYA-----QGADSVPEMPER 507
QY 577 ATKNEYAFYKNCITLNTGTGIRFEIATEAKNMFELPDKEQKVLMTIISKRLQNLAVGFVK 636
DB HLKNTYA-----GLQVNVVILPGKTPVVAEVK 534
QY 637 HYCDHTIGVANHITSETVTTKALASLRHEKSGKRIFYQIALKINAKLGGINQELDWSEIA 696
DB RVGDTVLGNATQCVQMKNV-----QRTTPQTLNLCIKINVKLGGVNNIL----- 579
QY 697 EISPEKERRKMTPLTMVGVGDVTHPTYSYSGIDYSIAAVVASINPGGTYIRNMIVTQEBC 756
DB -----LPQGRPPVFOQPV-IFLGADVTHPPAGDGKKPSIAAVGVGSMADHPNRYCATVRVQ-- 634
QY 757 RPEGRAVAGRETDILEAKFVKLLREFAENNDRAPAHIVVYRGVDSSEMLRVSHDEL 816
DB -----HROEIIQDLAAMVRELLIQFYKSTRFK-PTRIIFYRGVSEGGQCVLUHHEL 685
QY 817 RSLKSEVKQFMSERDEGPEPKYTFIVIQKRHNTRLLRRMEKDKPVNKNDLTPAETDVAV 876
DB LAIREACIKL-----EKDYQPGITIVVQKRHHTLTF-CTDKNERVKGSGNIPA----- 733
QY 877 AAVQWEEBDMKESKETGIVNPSGGTVDKLVSKYKDFDFLASHHGVLGTSRPGHYTMY 936
DB -----GTTVDTKITHPTFDFYLCSHAGIQGTSRPSHYHVLW 770
QY 937 DDKMGSDQBYVMYKMTYGLAFLSARCKKPSLPPVHVHLSCEKAKELYRKYEH 990
DB -----YHLVDKEH 823

ADJ94698
 ID ADJ94698 standard; protein; 860 AA.
 XX
 AC ADJ94698;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human eukaryotic translation initiation factor 2C2 (eIF2C2) protein.
 XX
 KW cytostatic; antiinflammatory; virucide; immunosuppressive; tumour;
 KW inflammatory; infectious disease; viral infection; degenerative;
 KW autoimmune; gene therapy; Argonaute family;
 KW eukaryotic translation initiation factor 2C2; eIF2C2; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004007718-A2.
 XX
 PD 22-JAN-2004.
 XX
 PF 10-JUL-2003; 2003WO-EP007516.
 XX
 PR 10-JUL-2002; 2002EP-00015532.
 PR 23-AUG-2002; 2002EP-00018906.
 XX
 (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA
 XX Tuschl T, Martinez J, Patkaniowska A, Urlaub H, Luehrmann R;
 XX WPI; 2004-122948/12.
 DR N-PSDB; ADJ94704.
 XX
 PT New single-stranded RNA molecule having a length from 14-50 nucleotides,
 PT useful for preventing or treating tumor, inflammatory, infectious, e.g.
 PT viral infections, degenerative and autoimmune diseases.
 XX
 PS Example; Fig 14; 73pp; English.
 XX
 CC The invention relates to a novel single-stranded RNA molecule having a
 CC length from 14-50 nucleotides where at least 14-20 of the 5'-most
 CC nucleotides are substantially complementary to a target transcript. The
 CC RNA molecule of the invention demonstrates cytostatic, antiinflammatory,
 CC virucide and immunosuppressive activities and may be useful for
 CC inhibiting the expression of a target gene in vitro or in vivo,
 CC preferably for preventing or treating diseases associated with the
 CC overexpression of at least one target transcript. The diseases may be
 CC selected from tumour diseases, inflammatory diseases, infectious diseases
 CC such as viral infections, degenerative diseases and autoimmune diseases.
 CC Furthermore, the molecules of the invention may be utilised during gene
 CC therapy. The current sequence is that of the human eukaryotic translation
 CC initiation factor 2C2 (eIF2C2) protein of the invention.
 XX
 SQ Sequence 860 AA;

Query Match 10.6%; Score 566; DB 8; Length 860;
 Best Local Similarity 23.1%; Pred. No. 2.7e-38;
 Matches 234; Conservative 160; Mismatches 370; Indels 250; Gaps 39;

QY 19 PEMKWLARPTGKCDGKFKYKVKVLLVWPKFSKIYDREYVEYVMTKEVLNRKPKPF 78
 DB 18 PIQGYAFKPPRPDPFGTSGRTIKLQANFFEMDIPKIDIVHYELDIKPEK-----C 67
 QY 79 PKTETPIEDRAKLFWQHLRHEKKQTDFILEDY--VFDEKDTVYSVRLNTVTSKMLVSE 136
 DB 68 PRVNRREIVEHM-----VQHFKTQ---IFGDKPVPDGRKNLYTAMPLPIGRDKV---E 115
 QY 137 KVVVKDSEKKDEKLEKKILYLTMLYRKPHLNFSENPEKDEEANSRYKFLKNVMTQ- 195
 DB 116 LEVTLPEGEKDR--IFKVSIKWVSCVSLQALHDHALSGRLPSVPFE---TIQALDVVMRHL 170
 QY 196 -KVRYPAPFVNEERIKVQFAKNFVVDNNSILRVPEFHPNRFQOSLEVAPRIEAWFGIYIG 254
 DB 171 PSMRYTPV-----GRSFTASEGCSNP-----LGGREWFVGFHQS 206

QY 255 IKELFDGEPVLNFAIVDKLFPYNAPKMSLLDYLLIVDPQSCNDDVRKDLKTKLMAGKMTI 314
 DB 207 VRPSL-WKMLNIDVSATAFYKA--QPVI8FVCEVLDFKSEEQ-----QKPL 251
 QY 315 RQARPPRIQLLENLKLK---CAEVDN-----EMSLRTERHLTFLDICEENSLVYKVTGK 367
 DB 252 TDSQRVFTKEIKGLKVEITHCCQMKRKRYVCNVTRRPASHQTF-PLQEGSGQVECT-- 308
 QY 368 SDRGRNAKYDITTLFLKIYENKKEPIEPHPLPLVKVKSAGKAYAVPMHEHLEVHEKPORYN 427
 DB 309 -----VAQY----FK--DRHKLVLRYPLPCLOVGOEQKHTYLPLEVNCNVAQORCIKK 356
 QY 428 RIDLVMQDKFLKRAI-RKPHDYKENTLMLKELDFSSSEELNFVERFLGCSKLOMIECPGK 486
 DB 357 LTD--NQTSMIRATARSAPDROBEISKMRSA5FNTDP--YVREFGIMVKDEMTDVTGR 412
 QY 487 VLKEPMLVNSNEQIKMTPV--IRGPOEKQLNV-----VPEKELCCAVFVNE 532
 DB 413 VLQPPSILYGRNKAIAITPVQGVMDMRNKQFHTGIEIKVWAIACFAPQO-CTEHLKSF 471
 QY 533 T-----AGNPCLEENDVVKFVTELIGGCKFRGIRIGANENRGAOSI--MYD 576
 DB 472 TEQLRKISRDAGMPIQOPCF-----CKYA-----QGADSVEMFR 507
 QY 577 ATKNEYAFYKNCNTLNTGIRFEIAATEAKMFERLPDEKQKLMFIIISKQLNAYGVFK 636
 DB 508 HLKNTVA-----GLQLVVVILPGKTPVYAEVK 534
 QY 637 HYCDHTIGVANOHITSETVTKALASLRHEKSGKRIFQIAIKINAKLGGINQELDNSEIA 696
 DB 535 RVGDTVLGMATQCVQKQNV-----QRTTPQTLNCLCKINVKLGGVNNIL----- 579
 QY 697 EISPEEKERRKMTPLWYGVIDVTHPTSYSGIDYSIAAVVASINPGGTIYRNMIIVTQEC 756
 DB 580 --LPQGRPPVPQPV-IFLGADVTHPPAGDKKPSIAAUVGSDMAHPNRCAIVRQQ-- 634
 QY 757 RQGERAVAHGRERTDILEAKFKVLLREFAENNDNRAPAHIVVYRDGVSDEMLRVSHDEL 816
 DB 635 -----HRQETIQDLAAMVRELLIQFYKSTRFK-PTRIIFVRDGVSEGFQFQVLHHEL 685
 QY 817 RSLKSEVKQFMSRDEGDEPEPKYTFIVIQGHNTLLRMKEXKPKVNVKDLTPAETDVAV 876
 DB 686 LAIREACIKL-----EKDYQPGITFIVQGRHTRLF-CTDKNERVKSNGNIPA----- 733
 QY 877 AAVKQWEEDMKESKETGINVPSGTTVDKLVSKYKFPDPLASHHGVLTGSRPGHYTVWY 936
 DB 734 -----GTTVDYKITHPTFDFYLCSHAGIQGTSRFSHYVLW 770
 QY 937 DDKGMSQDEVYKMTYGLAFLSARCKRPI8LPPVPHYAHLSCEKAKELYRTYKEH 990
 DB 771 DDNRFSDBELQILTYQLCHTYVYVTRSVSIPAPAYVAHLVAFRAR-YHLVDKEH 823

RESULT 15
 ABB67046
 ID ABB67046 standard; protein; 950 AA.
 XX
 AC ABB67046;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 27930.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers BW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL11149.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 27930; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 950 AA;

Query Match 10.5%; Score 560.5; DB 4; Length 950;
Best Local Similarity 22.9%; Pred. No. 9.2e-38;
Matches 225; Conservative 170; Mismatches 386; Indels 201; Gaps 35;

QY 38 KVVLLVNNFKSSKIYDREY-YEYEVKMTKVLNRKPKPKPKTEIPIDRAKLFQW 96
DB 112 RPVLKRNHFQVT---MPRGYVHHYDINIQDKPRKNREI---IETVHAYSIFG-- 163
QY 97 LRHEKKQTDFILEDYVFDEKDTVYSVCRNLNTVTSMKLVSEKVKVKKDSEKKDKLEKKIL 156
DB 164 -----VLKP-VFDGRNNLYTRDPLPIGNERL---ELEVTLPGEGKDR----- 201
QY 157 YTMILTYRKFFHLNFSRENPEDDEANRSYKFLKNVMTOKRYAPVPVNEEIKVQAKNFV 216
DB 202 ---IFRVTIKWAQVSLFLEALEGR-----TRQIPYDAILDVVMRHLPSMT 248
QY 217 YD--NNSILRVPEPHDPNRFQSLVAPRIEAWFGIYIGIKELPDGEPVLFNFAIVDKLF 274
DB 249 YTPVGRSPFSSPEGYYHP-----LGGREWFQFQHSVRP-SQWKMLNIDVSATAF 299
QY 275 YNAPKMSLLDYLLIYDPOSNDVVRKDKLTKLMAGKMTIRQAARPRIRQLLENKLCKA 334
DB 300 YKA--QPVIDFMCVLDIRDINEQ-RKPL-----TDSQRVKFTKEIKGLKIEIT 345
QY 335 EVDNEMSLRTERHLTFLDCEENSLVYKVTGSDRG-----RNAKYDTITLFXI 384
DB 346 H-----CGQMRKRYKVCNVTRPAQMQSPFLQLENGQTVECTVAKY 386
QY 385 Y-BENKKFTEPHPLPVKSGAKEYAVPMHELVHEKPKQRYKNRIDLVQDKFLKRAFR 443
DB 387 FLDKYRMKLYRPHLPCLQVQGEHKHTYLPLEVNI-VAGQRCIKKLTDMQSTMIKATAR 445
QY 444 KPHDYKENTLKMKELDSSSELNFVERFGLCSKLQMECPGKVLKBPMLVNSVNEIQIK 503
DB 446 SAPDREREINLVKRADFNND--SYVQEFGLTITSMSMEVGRVLPPLPKLYGGR----- 498
QY 504 TPVIRGFQKQLNVVPEKELCCAVFVNNETAGNPLEENDVVKFYTELIGGCKFRGIRIG 563
DB 499 --VSTGLTGCQQL-FPPQNKVSLA-----SPNQGVMDMRGKQFFTG--EIRIW 541
QY 564 ANENRGAQ-SIMYDATKNEYAPYKNCNTLNTG---IG-----RFEIATAEAKMFERLPDK 614

Search completed: July 5, 2006, 13:28:31
Job time : 206 secs

DB 542 AIACFAPQRTVREDALRNFTQLOKISNDAGNPIIGPCFCYKATGPDQVPMFRYLKIT 601
QY 615 EOKVLMFIIISKROLNAYGFVKHYCDHTIGVANQHITSETVTYKALASLRHEKSGKRIFYQ 674
DB 602 FPLQLVAVVVLPGKTPVYAEVKRVGDTVLGMATQCVAQNVNKT-----SPOTLSN 652
QY 675 IALKINAKLGGINQELDWSAEIAEISPEEKERKTPMLTMYVGIDVTHPTSYSGIDYSIAA 734
DB 653 LCKINVKLGGINSIL---VPSIRPKVFNEP-----VIFLGADVTHPPAGDNKKPSIAA 703
QY 735 VVASINPGGTIYRNMIVTQECCRGERAVAHGRERTDILEAKFVKLLREFAENNDNRAPA 794
DB 704 VVGSMDAHPSRVAAATVRVQQ-----HQEIIQELSSVMRELLIMFYKSTGGYKPH 753
QY 795 HIVYRDGVSDSEMLRVSHDELRSLSKSEVKQPMSESDGEDPE--PKYTFIVIOKRNTRL 852
DB 754 RIILYRDGVSEGGFPHVLQHELTAIREACIKL-----EPEYRPGITFIVVQKRHTRL 806
QY 853 LRMEKDKPVVNKDLTPAETDVAAVAQWEEDEKSKETGIVNPSSTGTTVDKLIIVSKYK 912
DB 807 FCAEKKK-----QSGKSG--NIPAGTIVDVGITHPT 836
QY 913 FDFFLASHHGVLGTSRPGHYTVMYDDKGMSEDEVYKMTYGLAFLSARCKRPISLPVPHY 972
DB 837 FDFYLCSHQGIQTSRPSHYHVLMDNHNHFDSDQLCLTVQLCHTVYVRCRSVSIPAPAY 896
QY 973 AHLSCAKELYRTTYKEHYIGD 994
DB 897 AHLVAFRAR-VHLVEKEHDSGE 917

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM protein - protein search, using sw model

Run on: July 5, 2006, 13:28:51 ; Search time 48 Seconds
(without alignments)
2044.606 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 5349

Sequence: 1 MSNPFPELEKGFYRHSLEDPF.....RHEMEHFLQTNVYKPGMSFA 1020

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4677.5	87.4	905	2 T23510	hypothetical prote
2	604.5	11.3	997	2 A84578	Argonaute (AGO1)-1
3	557	10.4	1000	2 T22391	hypothetical prote
4	556	10.4	813	2 JC6569	translation initia
5	546.5	10.2	1032	2 T23164	hypothetical prote
6	546.5	10.2	1035	2 T23165	hypothetical prote
7	540.5	10.1	891	2 T32079	hypothetical prote
8	537	10.0	958	2 S41013	hypothetical prote
9	537	10.0	1040	2 D88568	protein ZK757.3 (i
10	510.5	9.5	1014	2 H86438	protein T19823.7 (i
11	488	9.1	988	2 T52134	Zwille protein [im
12	484	9.0	834	2 T41568	argonaute-like pro
13	473	8.8	898	2 T51564	hypothetical prote
14	428.5	8.0	930	2 A84568	Argonaute (AGO1)-1
15	402	7.5	887	2 T01113	translation initia
16	356	6.7	185	2 T27784	hypothetical prote
17	347.5	6.5	1139	2 T33275	hypothetical prote
18	311.5	5.8	674	2 T19268	hypothetical prote
19	311.5	5.8	1030	2 T16114	hypothetical prote
20	309.5	5.8	722	2 T30995	hypothetical prote
21	302.5	5.7	990	2 F88925	protein T22H9.3 (i
22	296.5	5.5	965	2 T22933	hypothetical prote
23	294	5.5	1032	2 T19324	hypothetical prote
24	287.5	5.4	824	2 T20351	hypothetical prote
25	272	5.1	1057	2 T16676	hypothetical prote
26	259.5	4.9	939	2 T18974	hypothetical prote
27	247.5	4.6	881	2 T31818	hypothetical prote
28	237	4.4	945	2 T23965	hypothetical prote
29	223	4.2	892	2 T15195	hypothetical prote

30	211.5	4.0	944	2 T18911	hypothetical prote
31	206.5	3.9	697	2 T15179	hypothetical prote
32	192.5	3.6	2269	2 T28677	rhoptory protein -
33	190.5	3.6	946	2 S28061	SCP1 protein - rat
34	183	3.4	363	2 T34486	hypothetical prote
35	182.5	3.4	587	2 T34339	hypothetical prote
36	175.5	3.3	1979	2 C71622	hypothetical prote
37	175	3.3	2253	2 T30336	nuclear/mitotic ap
38	175	3.3	2401	2 T28676	rhoptory protein -
39	174.5	3.3	1302	1 JC6009	surface-located me
40	172.5	3.2	1660	2 A84647	hypothetical prote
41	172	3.2	1939	2 T18372	repeat organellar
42	170.5	3.2	324	2 G87774	protein C24A11.3 (
43	170	3.2	1170	2 A72287	hypothetical prote
44	169.5	3.2	2829	2 A42771	reticulocyte-bind
45	168.5	3.2	980	2 E71606	hypothetical prote

ALIGNMENTS

RESULT 1

T23510

hypothetical protein K08H10.7 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T23510

R;Gardner, A.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19750

A;Accession: T23510

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-905 <WIL>

A;Cross-references: UNIPROT:Q9XU82; UNIPARC:UPI000017BAA9; EMBL:Z83113; PIDN:CAB05546.1

A;Experimental source: clone K08H10

C;Genetics:

A;Gene: CESP:K08H10.7

A;Map position: 5

A;Introns: 19/3; 86/2; 190/3; 209/2; 269/1; 341/2; 468/3; 671/3; 832/3

Query Match 87.4%; Score 4677.5; DB 2; Length 905;
Best Local Similarity 88.7%; Pred. No. 1e-254;
Matches 905; Conservative 0; Mismatches 0; Indels 115; Gaps 1;

QY 1 MSNPFPELEKGFYRHSLEDPDKCKYKLLVNNVFPKSSKIYDREYVE 60

DB 1 MSNPFPELEKGFYRHSLEDPDKCKYKLLVNNVFPKSSKIYDREYVE 60

QY 61 YEYKMTKEVLRNKPGRPPKTEIPIPDRAKLFQWHLRHEKKQTDPILEDYVFEKDTVY 120

DB 61 YEYKMTKEVLRNKPGRPPKTEIPIPDRAKLFQWHLRHEKKQTDPILEDYVFEKDTVY 120

QY 121 SVCLNTVTSKMLVSEKVKVKKSEKDEKLEKKILYTMILTYRKFFHLNFSRENPKDE 180

DB 121 SVCLNTVTSKMLVSEKVKVKKSEKDEKLEKKILYTMILTYRKFFHLNFSRENPKDE 180

QY 181 EARNRSYKFLKNVNTQKRVAPFVNEBIKVQFANFVYDNNISILRVPSFHPNRFQESLE 240

DB 181 EARNRSYKFLKNVNTQKRVAPFVNEBIKVQFANFVYDNNISILRVPSFHPNRFQESLE 240

QY 241 VAPRIEAMFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLIIVDPQSCNDDVR 300

DB 241 VAPRIEAMFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLIIVDPQSCNDDVR 300

QY 301 KDLTKLMAGKMTIQAARPRITROLLENKLKCAEYVNDNEMSLRTERHLTFLDLCHNSL 360

DB 301 KDLTKLMAGKMTIQAARPRITROLLENKLKCAEYVNDNEMSLRTERHLTFLDLCHNSL 360

QY 361 VYKVTGKSDRGNAKKYDTTLTKIYEENKFTIEFPHLPLVKVKSAGKAYAVPMHELVHE 420

DB 361 VYKVTGKSDRGNAKKYDTTLTKIYEENKFTIEFPHLPLVKVKSAGKAYAVPMHELVHE 420

QY 421 KFORYNKIDLVQDKFLKRAATKPHDYKENTLKMELDSSSEELNFVERFLGSKLQM 480
 DB |||||
 QY 421 KFORYNKIDLVQDKFLKRAATKPHDYKENTLKMELDSSSEELNF 468
 DB |||||
 QY 481 IBCPGKVLKEPMLVNSVNEQIKMTPIVIRGQKQLNVVPEKELCCAVFVNNETAGNPCLE 540
 DB |||||
 QY 469 ----- 468
 QY 541 ENDDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIRPEITA 600
 DB |||||
 QY 469 ----- 485
 QY 601 ATEAKNMPERLPDKEQKLMFIIISKROLNAYGFVKHYCDHTIGVANQHTSETVTKALA 660
 DB |||||
 QY 486 ATEAKNMPERLPDKEQKLMFIIISKROLNAYGFVKHYCDHTIGVANQHTSETVTKALA 545
 DB |||||
 QY 661 SURHEKSGKRIFYQIALKINAKLGGINQELDWSAEISPEKERRKTMPLTMYVGIDVT 720
 DB |||||
 QY 546 SURHEKSGKRIFYQIALKINAKLGGINQELDWSAEISPEKERRKTMPLTMYVGIDVT 605
 DB |||||
 QY 721 HPTSGIDYSIAAUVASINPGGTYIRNMIVTQECRPGERAVAGHRETDILEAKFKVL 780
 DB |||||
 QY 606 HPTSGIDYSIAAUVASINPGGTYIRNMIVTQECRPGERAVAGHRETDILEAKFKVL 665
 DB |||||
 QY 781 LREFANNDRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSVKQFMSEKEDGEDPEPKYT 840
 DB |||||
 QY 666 LREFANNDRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSVKQFMSEKEDGEDPEPKYT 725
 DB |||||
 QY 841 FVIOQRHNTLLRMEKDKPVVKNKDLTPAETDVAVAVKQWEEDMKSKSETGIVNPSGG 900
 DB |||||
 QY 726 FVIOQRHNTLLRMEKDKPVVKNKDLTPAETDVAVAVKQWEEDMKSKSETGIVNPSGG 785
 DB |||||
 QY 901 TVVDKLVSKYKFPDFLASHHGVLGTSRPGHYTVMYDDKGMQDEVYKMTYGLAFLSARC 960
 DB |||||
 QY 786 TVVDKLVSKYKFPDFLASHHGVLGTSRPGHYTVMYDDKGMQDEVYKMTYGLAFLSARC 845
 DB |||||
 QY 961 RKPISLPVPVHYAHLSCAKELRYKHYIGDYAQPRTREMEHFLOTNVKYFCMGSA 1020
 DB |||||
 QY 846 RKPISLPVPVHYAHLSCAKELRYKHYIGDYAQPRTREMEHFLOTNVKYFCMGSA 905
 DB |||||

RESULT 2

AQ4678
 Argonaute (AGO1)-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: AB4678
 R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487; PMID:10617197
 A:Accession: AB4678
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-997 <STO>
 A:Cross-references: UNIPROT:Q9SJK3; UNIPARC:UPI00001256F1; GB:AE002093; NID:94510428; PI
 C:Genetics:
 A:Gene: At2G27880
 A:Map position: 2
 C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 11.3%; Score 604.5; DB 2; Length 997;
 Best Local Similarity 23.8%; Pred. No. 4.5e-26;
 Matches 242; Conservative 164; Mismatches 353; Indels 259; Gaps 41;
 QY 38 KVVLLVNVKFPSSKIYDREYVEYKMTKEVLNRPKPGKPPKTEIPIIDRAKLFWQHL 97
 DB |||||
 QY 158 KKVMTAN--HPLVQVARDLYHYDVSINPEVISK----- 190
 DB |||||
 QY 98 RHEKQTDFILEDDYVDEKOTVYVCRNLNTVTSKMLSEKVVKKDSE---KKDEKDLBK 154
 DB |||||

DB |||||
 QY 191 -----TVNRNVKMLLVKN---YKDSHLGKSPAYD-GRK 220
 QY 155 ILYTM--ILTVRKKTHLAPSRNPENPEKDEANRSYK-F-LKV------MTQKRVYA 200
 DB |||||
 QY 221 SLYTAPLPFDOSKEFVNVLAEKRADGSGGKDPFKVAVKNNVTSTDLVYQQLDLDRKOREA 280
 DB |||||
 QY 201 PPNVEIIVQPAKPNFYVDNNSILRVPESPHDNRRFEQSL-----EVAPRIEAWFGI 251
 DB |||||
 QY 281 PY--DTIQVLVDVLVDKFSNDYVSGRSF-----FHTSLGKADRGRLGDLGIEWRGY 333
 DB |||||
 QY 252 YIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLLVDPQSCNDDVRKDKLTKLMAGK 311
 DB |||||
 QY 334 PQSLA-LAQMLGLSINIDVSARSFY-----EPIVWTFDFISKELNIRDL--N 375
 DB |||||
 QY 312 MTIROAARPRITQLLENLKLKCAEYVDNEMSRLETHLTFDLCEENSIVYKVTGKSD-- 369
 DB |||||
 QY 376 RPLRSDRLKVKVKVRLTKVLLH-WN-----CTKSA---KISGISSLP 415
 DB |||||
 QY 370 ---RGRNAKYDYDTLTKIYEENKCF-IEFPHLPVVKVSGAKAYAVPMHELVHEKPOR 424
 DB |||||
 QY 416 IRELFTLEDKSEKTVQVFAEKYNYRVKYQALPAIQGSDTRPVVLPNELCOIDE-QOR 474
 DB |||||
 QY 425 YKNRIDLVMDQKFLKRAATKPHDYKENTLKMELDSSSEELNFVERFLGSKLQMIECP 484
 DB |||||
 QY 475 YTKRLNEKQVTTALLKATCORPPD-RENSIGNLVKNYNDLDS--KEFGMSVTTQLASIE 531
 DB |||||
 QY 485 GKVLKEPMLVNSVNEQIKMTPIVIRG-----FOEKQLNVVPEKELCCAVFVNNETAGNPCLE 540
 DB |||||
 QY 532 ARVLPPPMKLVKHSKERNVPRLGQNMIDKQVNGAKVTSWTCVSFSTRIDRGLP--- 588
 DB |||||
 QY 541 ENDDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIRPEITA 600
 DB |||||
 QY 589 ---QEFCKQLIGMCSKGM-----EFKQPAIPIFISCPPEH-----IEEA 625
 DB |||||
 QY 601 ATEAKNMPERLPDKEQKLMFIIISKROLNAYGFVKHYCDHTIGVANQHTSETVTKALA 660
 DB |||||
 QY 636 LLD---IHKRAPGLQ---LLIVLPDVTGYSYKIKRICETELGIVSQCCQPRQVKNL-- 676
 DB |||||
 QY 661 SLRHEKSGKRIFYQIALKINAKLGGINQELDWSAEISPEKERRKTMPL-----TMIV 715
 DB |||||
 QY 677 -----NKQYMNENVALKINVTGGRNTVLN-----DAIRRNIPILTDPTIIM 718
 DB |||||
 QY 716 GIDVTHPTSYSGIDYSIAAUVASIN-PGGTYIRNMIVTQ---EE-----CRPGERAV 763
 DB |||||
 QY 719 GADVTHPQPGEDSSPSIAAUVASMDWPEINKYRGVLSAQAHREELIQDLYKLVQDPQRL 778
 DB |||||
 QY 764 AHGRETDILEAKFVKLLREFANNDRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSEV 823
 DB |||||
 QY 779 VH-----SGLIREHETAFRA-----TGQIPORIIFYRDGVSGSQFSQVLLHEMTAIRKAC 829
 DB |||||
 QY 824 KQPMSEKEDGEDPEPKYTPIVIOQRHNTLLRMEKDKPVVKNKDLTPAETDVAVAVKQWE 883
 DB |||||
 QY 830 NSLO-----ENYVPRVTVIVQKRHHTLRFPEQHG-----NKDMT----- 864
 DB |||||
 QY 884 EDMKESKETGIVNPSGGTVDKLVSKYKFPDFLASHHGVLGTSRPGHYTVMYDDKGMQ 943
 DB |||||
 QY 865 -----DKSGNIQP--GTVDVTKICHPNEFDYLNSHAGIQCTSRPAHVHVLDDENGFTA 916
 DB |||||
 QY 944 DEVYKMTYGLAFLSARCCKPISLPVPVHYAHLSCAKELRYKHYIGDYAQPRT 1001
 DB |||||
 QY 917 DQLQMLTNLCYTVARCTKSVIPPPYYAHLAAPRA-----RYNMESEMSDGGSSRSR 970
 DB |||||

RESULT 3

T22391
 hypothetical protein F487.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T22391; T23234
 R.Coates, L.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19559
 A:Accession: T22391

[illegible]

```

Db          910 VLWDDNLLTADELQQLTYQMCHTYVRCRSVSIPAPAYAHVLVAPRAR-YHLVDRHDSG 968
QY          994 DYAOQ 998
Db          969 EGSQP 973

RESULT 4
JC6569
translation initiation factor eIF-2C - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: PC6505; JC6569
R:Zou, C.; Zhang, Z.; Wu, S.; Osterman, J.C.
Gene 211, 187-194, 1998
A:Title: Molecular cloning and characterization of a rabbit eIF2C protein.
A:Reference number: JC6569; MUID:98267198; PMID:9602122
A:Accession: PC6505
A:Molecule type: protein
A:Residues: 336-359;380-409;694-711 <ZOU1>
A:Cross-references: UNIPROT:O77503; UNIPARC:UPI00001790F8; UNIPARC:UPI00001790F8
A:Experimental source: liver
A:Accession: JC6569
A:Molecule type: mRNA
A:Residues: 1-813 <ZOU2>
A:Cross-references: UNIPARC:UPI000012D07F; GB:AF005355; NID:g3253158; PIDN:AAC3
C:Superfamily: rabbit translation initiation factor eIF-2C
C:Keywords: liver

Query Match          10.4%; Score 556; DB 2; Length 813;
Best Local Similarity 23.6%; Pred. No. 1.8e-23;
Matches 230; Conservative 152; Mismatches 346; Indels 248; Gaps 40

QY 57 EYTYEYVMTKVEVLNRKPGKPPKTKTPIPIPDRAKLFWQHLRHKKQTFDILEDY-VVD 114
Db 7 DIHYE-----LDIKPEK-CPRVNRIVEHNM-----VQHFKAQ---IFGDRKPVVD 49

QY 115 EKDITVSVCLRINTVTSKMLVSEKVVKKDSKKDSKDEKLEKKILYTMILTYRKKFHLNFSRE 174
Db 50 GRKILYTMPLPIGREKV-----ELEVTLPGEKOR--IPKVSIRKVVSCVSLQALHDALSGR 104

QY 175 NPEKDEBANRSYKPLKQNVMTQ--KVRYAPFVNEIKVQFAKNFVYDNNLSILRVPESHDP 232
Db 105 LPSVPPF--TIQALDVVMHLSMRYPV-----GRSPFTASEGCSNP 145

QY 233 NRPEQSLEVPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLDYLLIVDP 292
Db 146 -----LGGREVVFGFHQSVRPSL-WKMLNIDVSATAFYKA--QPVIEFVCEVLDF 194

QY 293 QSCNDDVVRKDKTKLMAGKMTIROAEPRIQLLENLKLK---CAEVDN-----EMSLRT 345
Db 195 KSIEEQ-----QKPLTDSQVKTKEIKGLKVBITHCGQMKRYKRVNCNTRRP 242

QY 346 ERHLTFLDLCENSLVYKTKSDRGKNAKKYDTTLTKIYEENKKTIEFPHLPVLVKVKG 405
Db 243 ASHQTF-PLQESQTVECT-----VAQY-----FK--DRHKLVLRYPHLPCLQVQE 287

QY 406 AKEYAVPMHELVHEKPORYKNRIDLVMDKFLKGRAT-RKPHDYKENTLQWLKELDFSS 464
Db 288 QKHTYLPLEVCNIVAGORCIKKLTD--NQTSMTIRATARSAPDQREISIKLMRSASFNTD 345

QY 465 ELNFVERFGLCSKLOMECEPKVLKEPMLVSNVNEQIKMTPV--IRGFQEKQLNV----- 517
Db 346 P--YVREFGIMVDEMDTVDVGRVLQPLSPSILYGRGNKAIAIPVQGVDMRNKQFHTGIEIK 403

QY 518 -----VPEKELCCAVFVNNET-----AGNPFLEENDVWVKFYTELIGCK 556
Db 404 VWAIACFAPORQ-CTEVLHKSFTLEQLRKISRDAGMPIQGPCP-----CK 447

QY 557 FRGIRIGANENRGAQSI--MYDATKNEYAFYQNTCLNTLTGTIGRFEIAATEAKNMFERLPDK 614
Db 448 YA-----OGADSVGPMFRILKNTYA----- 467

```


A:Cross-references: UNIPROT:Q21079; UNIPARC:UPI000008666E; EMBL:Z68750; PIDN:CAA92970.1.1
A:Experimental source: clone K01A6
F:Lennard, N.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19981
A:Accession: T25100
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1035 <WI2>
A:Cross-references: UNIPARC:UPI000008666E; EMBL:Z68300; PIDN:CAA92619.1; GSPDB:GN00022;
A:Experimental source: clone T22B3
C:Genetics:
A:Gene: CESP:T22B3.2b
A:Map position: 4
A:Introns: 19/1; 74/3; 132/2; 427/2; 499/3; 610/1; 669/1; 756/3; 798/2; 964/2; 1006/1
C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match	10.2%	Score 546.5	DB 2	Length 1035
Best Local Similarity	22.3%	Pred. No. 8.5e-23		
Matches 225	Conservative 175	Mismatches 397	Indels 213	Gaps 38
Qy	38	KVLLLVNWKFSKIVDREYVEYVMTKEVLNRKSGKPPKTEIPIPDRAKLFWQ--	95	
Db	95	RIPIVKSNFPAVDLKNPFWVQVTHVEV-----HFGCRKLDKDEMRI-----IFKAV	143	
Qy	96	-----HLRHEKKQTDFILEDYVDEKQTVISYCRLL-----NTVTSKMLVSEKVVKKSEKKDE	148	
Db	144	SDHENIFHNK-----FAL-----AYDGAHQLYTVARLEFPDQGSVRLDCEATI.PKDNRRDTR	196	
Qy	149	KOLEKKILYMLTYRKPHLNFSPENPEKDEEANRSYKFLNVMVTKRVYAPVPVNEBIK	208	
Db	197	CAISIQNVGPVLEEMOR-----TRTN-ILDERVLTPIQLIDITCROSLT-CPLLKN-----	245	
Qy	209	VQFAKNFVYDNNSILTRVPESPHDNRFEQSLSEVAPRIEAMFGIYIGIKELDCEPVLNFA	268	
Db	246	-----SANTYWKSSCYRIPTA-----AGQALDEGGKEMTGFSSAHIASNRYRPLNID	296	
Qy	269	IVDKLFPYNAKPMGLDLYLLIVDPQSCNDDVKDKTKLMAGKMTIRQAAARPRIQLLEN	328	
Db	297	VAHTAFYKT-RITVLQFMCVDLNBERTSKPNRNNPRGPCAPGVRGGRGARGGSYQNFN	355	
Qy	329	LKLKCAEVRWENMSRLTERHLTF-LDLCEE-----NSLVYKVTGKSDRGNAK-----	375	
Db	356	RGPPGANVRDD-----FGNGGLTFTMDLTSLRDTQLSSFETRIFGDSIRGMKIRATHRPNAI	411	
Qy	376	-----KYDTTLFK-IYEENKKPI-----BPPHPLPLVKVKSAGKEYA	410	
Db	412	RVYKVNLSQLPADKLMFQIGIDEGRQVVCVADYFSEKYGLPKYKPLPCLHVGPPTNIP	471	
Qy	411	VPEHLEVHEKPYKRNRIIDLWQDKFLKRAKTRKPHDYKENTLKMKLKELDFSSSEELNFVE	470	
Db	472	LPMEHCLLI-DSPOKYNKMTKEKOTSIAIKAAAVDATORBEDRIKQLAAQAASFGTDP--FLK	528	
Qy	471	RFLGCSKLQMIQCEGKVLKBPMLVNSVNEQIKWTFVIRGFOEKOLNVVPEKLCAPV--	528	
Db	529	EFGVAVSSQMIEFSARVIQPP-----PIMFGNKRNSINPV-----VFPK	567	
Qy	529	-----VYNETAGNP--CLEBNDVV-----KFYTELIGGCK-----FRGIRIGANERGAQSIM	574	
Db	568	DGWSMDHQTLYMPATCRSYSMIALVDPDRDTSLSQTFCSQLTKKATAMGNWFWPDPDLVK	627	
Qy	575	YDATQNEYAFYKNCNTLNTGIG-RPEIAATEAKWPFELPQEKVLMFLIISKRLNAYG	633	
Db	628	YGRSKEDV-----CTLTEADEYRVNTVCD-----CIIIVLOKNSDSIYM	669	
Qy	634	FVRHYCDHTTIGVANQHTTSETVTTKALASLRHEKSGSKRIFYQIALKINAKLGGINQELDWS	693	
Db	670	TVKEQSDIVHGIIMSQCVMKNVSRP-----TPATCANIILKNWKGGINSR-----	717	
Qy	694	EIAIESPEEKERKTMPLTMVGDIVTHPT-SYSGIDY-SIAAVVASINPGGTIYRNMIV	751	
Db	718	-----VADQITNKYLDVQPTVVGVDVTHPTQAEWRNMMPVAAIVANVDLLPOSYGANVK	773	

QY 752 TQEBRCGERAVAHGRERTDILEAKFVKLLREFAEENNDRAPAHIVVYRDGVSDSMLRV 811
 DB 774 VQKCR---ESVY---LLDAIRERIITFYR-----HTQKQPARIIIVYRDGVSEGFSEV 822
 QY 812 SHDELRSLSKSEVKQPMSEDEGDEPPKPTFVIQKHNTLLRRMEKOKPVVKNKULTPAE 871
 DB 823 LREBIQSIRTACLAI-----AEDFRPPTIYIVQKRHHARI FCKFPNDMVGKAKNVPP-- 875
 QY 872 TDVAVAQWQEEDEKSKETGIVNPSGGTTVDKLI VSKYKDFFLASHHGVLTGSRPGH 931
 DB 876 -----GTTVDGTGIVSGEGDFFLCSHYGVQGTSRPAR 907
 QY 932 YTVMYDDKMGSDQEVYKMTYGLAFLSARCRKPISLPVVHYAHLSECAK 981
 DB 908 YHVLDECKFTADEIQNITYGMCHTYGCTRSVSITPTVYVADLVATRAR 957

RESULT 7
 T32079
 hypothetical protein T07D3.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T32079
 C:Lamar, B.; Wamsley, P.; Twyman, B.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of *C. elegans* cosmid T07D3.
 A:Reference number: 221121
 A:Accession: T32079
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-891 <LAM>
 A:Cross-references: UNIPARC:UPT0000002DB7; EMBL:AF016682; PIDN:AAB66187.1; GSP:
 A:Experimental source: strain Bristol N2; clone T07D3
 C:Genetics:
 A:Gene: CESP:T07D3.7
 A:Map position: 2
 A:Introns: 34/2; 92/2; 206/1; 560/1; 879/3
 C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 10.1%; Score 540.5; DB 2; Length 891;
 Best Local Similarity 23.4%; Pred. No. 1.5e-22;
 Matches 197; Conservative 138; Mismatches 332; Indels 175; Gaps 26

QY 189 LKNVMTOKRVYAPVFNBEIKVQAKNFYDNNLSILR-----VPESFHPD----- 232
 DB 163 LDDAMEGRVQVP-----HEAVQSIDVLRHLPSLKYTPVGRSFPTPGVMKPG 211
 QY 233 NRFEQSLVAPRIEAWFGIYGIKELFDGEPTVNAIFAIDKLFYNAPKMSLLDYLLLVDP 292
 DB 212 MQMHQESKLGGRVWFEGFHQSVRP-SQWKMLNIDVSATFYRA--MPVIEFVAEVL-- 266
 QY 293 QSCNDVDVRKDLTKYLMAGKMTIROAAPRIOLLENLKLCAEVDNEMSLRTEHLTFL 352
 DB 267 -----ELFVQALAEALSDAQVRKFTKEIRGLKIEITH----- 300
 QY 353 DLCEENSLVYKTKSGDRGNAKKY-----DTTLPKLYBENKKF-IEFPHLPVK 401
 DB 301 --CGAVRKRYVCNVTTRPAQTQTFPLQLENGQTIECTVAKYFPDKYRIQLKYPHLPCLQ 358
 QY 402 VKGAKYAVVPMHELVHEKPPQRYKNRIDLVQMDKFLKGRATRKPHDYKENTLKMKEIDF 461
 DB 359 VGEQKHLYLPPEVCDI-VPGQRCLKGLTDVQTSMTIKATARSAPEREREICKLVSKAEL 417
 QY 462 SSELBNFVERFGLSKLQWLECKVKLKEPMLVNSVNEQIKQTPVIRG---PQKQLNVV 518
 DB 418 SADP--FAHBFGITINPAMTEVKGRLVSPKLLYGGRRHATTALPNQGVMDMRGKQFHTG 475
 QY 519 PE-KELCCAVFVWNETAGNCLBENDVVVFYTELIGGCKFRGIRIGANENRGAOSIMYDA 577
 DB 476 MEVRTWAIACF-----AQOSHVKENDLRMFTTQL-----QRISTDA 511
 QY 578 TKNEYAPYKNTLNTWIGRFEIATEAKNMPERLPDKEQKVLMIIFIIISKQLNAYGPKVH 637

```
Db 512 GMPILCTPMEKCYASGEQVE-----PMFKYLKQTSYSAIQILIVVLPCKTPIYAEVKR 564
Qy 638 YCDHTIGVANQHITSETVTYKALASLRHEKSGKRIFYQIAIKINAKLNGGINQBLDWEISIAE 697
Db 565 VGDVTLGIATQCVOAKNAIRT-----TPQTLNLCCLKMNVKLGGVNSIL----- 608
Qy 698 ISPEEKERKTMPLTMYGIDVTHPTSYSGIDYSYAAVVASINPGGTIYRNIMVTOEBCR 757
Db 609 -LPNVRPRIFNBEV-IFLGCDDTHPAAGDTRKPSIAAVGVGSDMAHPSRYAATVRVQQ----- 663
Qy 758 PGERAVAHGRERTDILEAKFVKLLREFAEFNNDNRPAHIVVYRDGVSDSEMLRVSHDEL 817
Db 664 -----HQEIIITDLYMVRELLVQFYENTRFK-PARIVVYRDGVSEGLFNVLQYELR 715
Qy 818 SUKSEVKQPMSESDGDPKPKTYFIVIQKHNRLLRRMEKDKPVVKNKDLTPAETDVAVA 877
Db 716 AIREACVMLES-----GYQGITFIYAVQKRHHTRLFAADKADQ--VGK----- 756
Qy 878 AVKQWEEDMKESKETGINVPSSGTTVDKLIIVSKYKDFDFLASHHGVLGTSRPHYTMVD 937
Db 757 -----AFNIPGGTIVDVGITHPTPEDFFLCSHAGIQGTSRPSHYHLWMD 800
Qy 938 DKGMSODEYKMYTIGLAFLSARCKPISLPVPVHVVAHLSCERAKELYRTYKSHYIGDY-A 996
Db 801 DNDLTADELQQLTYQWCHTYVCTRVSISIPAPAYIAHLVAFRAR-YHLVDRDRHSGSEGS 859
Qy 997 QP 998
Db 860 QP 861

RESULT 8
S41013
hypothetical protein ZK757.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Dec-2004
C:Accession: S41013
R:Thomas, K.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41011
A:Accession: S41013
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-958 <THO>
A:Cross-references: UNIPROT:Q9TW94; UNIPARC:UPI0000179102; EMBL:Z29121
A:Genotons: 52/2; 350/2; 422/3; 592/1; 679/3; 721/2; 887/2; 929/1

Query Match 10.0%; Score 537; DB 2: Length 958;
Best Local Similarity 21.6%; Pred. No. 2.6e-22;
Matches 222; Conservative 170; Mismatches 375; Indels 262; Gaps 39;

Qy 45 NMFKFSSKIYDREYIEYEMTKVELNRRKPKGPPFKKTEIPIDRAKLFWQ-----HLR 98
Db 22 NFFAMDCLKPKMVVIQYHVEI-----HHPGCRKLDKDEMRI-----IFWKAVSDDHPNIF 70
Qy 99 HEKKQTDFILEYDFDEKDTVTVSVCRL---NTVTSKMLVSEKVKVKDKSEKDEKLEKXI 155
Db 71 HNK-----FAL---AYDGAHQLYTVARLEFPDPDQGSVRLDCEASLPKDNDRTRCAISIQN 123
Qy 156 LVTMILTYYKKFHLNFRSRENPEKDEANRSYKFLKNVMTQKVRYAPFVNEEIKVQFAKNF 215
Db 124 VGPVLEMQR-----TRTN-NLDERVLTPIQILDICQSIT-CPLLKN-----SANF 169
Qy 216 VYDNNILRVPSFHDPNRFEQSLEVAIPRIEAWFGIYIGIKELFDGEPVINFALDKFLY 275
Db 170 YTMKSSCYRIPTA-----AGQALDLEGGKEMWTGPFSSAHIASNYRPLNLDVAHTAFY 223
Qy 276 NAPKMSLLDYLIIYD----- 293
Db 224 KT-RITVLOFMCVDVLNERTSKPNRNNRPGGPGGPGGYRGGRGGSYGNFGNRGPP 282
Qy 294 SCN-----DDVRKD-----LTKTLMAGK---MTIRQAARPRIQLLENLKL 331
```

```
Db 283 GANVDDDFGGNGLTFTMTLSRDTQLSSFEETRIFGDAIRGMKIRAAHRP----- 331
Qy 332 KCAEYVMDNEMSRSLTERHLTFLDLCEBNSLVYKVGKSDRGRNAKKYDITTLFLKIYEENKKF 391
Db 332 NAIRVYKNSLQLPADKLMFQGIDEGRQV--VCSVAD-----YFSEKYGP 375
Qy 392 TEPFHLPLVKKVSGAKAYAVPMEHLEVHEKPORYKNRIDLVMQDKFLKQATRKPHDYKEN 451
Db 376 LKYPKLPCLHVGPPTRNIFLPMHECLI-DSPOKYNKMSEKQTSIIKAAAVDATOREDR 434
Qy 452 TLKMLKELDFSEELNFVERFGLCSKLOMIECPGKVLKEPMLVNSNEQIKMTPVIRGFO 511
Db 435 IKQLAAQASFGTDP--FLKEFGVAVSSQMIQTARVIQPP-----PIFNGGN 479
Qy 512 EKQLNVVPEKELCCAVF-----VYNETAGNP--CLEENDVY---KPYTELIGCK-- 556
Db 480 NRSVNPV-----VFPKDGSWTMDNOTLYMPATCRSYSMIALVDPDQTSLQTFQCSL 531
Qy 557 -FRGIRIGANENRGAQSIWYDATKNEYAFYKNCNTLNTGIG-RFEIAATEAKNMFERLPDK 614
Db 532 TMKATAMGMNFPWPDLVKYGRSKEDV-----CTLFTEIADBYRVNTVCD----- 577
Qy 615 EQKVLMFIIISKRQLNAYGVFVKHYCDHTIGVANQHITSETVTKALASLRHEKSGKRIFYQ 674
Db 578 ----CIIIVLQSKNSDIYTWVKEQSDIVHGINSQCVLKMNVRP-----TPATCAN 624
Qy 675 TALKINAKLGGINQBLDWSIAEISPEEKERRKTWPLTMVYVIGIDVTHPT-SYSGIDY-SI 732
Db 625 IVLKLMKMGGINRI-----VADKITNKYLVDPQTMVVGIDVTHPTQAEHRMNMPSV 677
Qy 733 AAVVASINPGGTIYRNIMVTQECPPGERAVAHGERTDILEAKFVKLLREFAEFNNDNR 792
Db 678 AAIYANVDLLPOSYGANVKKVQKCR---ESVYV---LLDAIRERITFYR-----HTKQK 726
Qy 793 PAHIVVYRDGVSDSMLRVSHDELRLSKSEVKQFMSEKDEGEPEPKYTFPIVQKHNRRL 852
Db 727 PARIIVYRDGVSEGFSEVLRREEIQSIRTACLAI-----AEDFRPPIYIVVQKRRHARI 781
Qy 853 LRRMEKDKPVVKNKDLTPAETDVAVAAVQWDEEDMKESKETGIVNPSSGTTVDKLIIVSKYK 912
Db 782 FCKYQNDMVYKAKNVPP----- 811
Qy 913 PDFFLASHHGVLGTSRPGHYTMVDDKGMQSDQEVYKMTYGLAFLSARCKRPSLSPVPVHY 972
Db 812 PDFYLCYSHYGVQGSTSRPARYHVLDECKFTEADISQIYGMCHTYGCRTRSISIPVYV 871
Qy 973 AHLSCSEKAK 981
Db 872 ADLVATRAR 880

RESULT 9
D88568
protein ZK757.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D88568
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/sec/C_elegans/ and www.sanger.ac.uk/projects/C.el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: D88568
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1040 <STO>
A:Cross-references: UNIPROT:P34681; UNIPARC:UPI000013BC28; GB:chr_III; PIDN:CAA82941.1;
C:Genotons:
A:Gene: ZK757.3
A:Map position: 3
C:Superfamily: rabbit translation initiation factor eIF-2C
```

Query Match	10.0%;	Score 537;	DB 2;	Length 1040;
Beat Local Similarity	21.6%;	Pred. No. 2.9e-22;		
Matches 222;	Conservative 170;	Mismatches 375;	Indels 262;	Gaps 39;
Qy	45	NWKFSSKIYDREYVEYVMTKEVLNRKPKGPPKKTETIPDRAKLFWQ-----HLR 98		
Db	104	NFFAMDLPKPKMVVYQYHVEI-----HHPGCRKLDKDEMRI-----IFWKA VSDHNPFI 152		
Qy	99	HEKQTDPILEDYVDEKDTVYSVCLR---NTVTSKMLVSKVKVKQSEKKDEKLEKKI 155		
Db	153	HNK-----PAL---AYDGAHQLYTVARLEFPDQGSVRLDCASLPKONDRTRCAISIQN 205		
Qy	156	LYTMILTVRKXPHLNFSPENPEKDEANRSYKFLKNVMTQKRVYAPFVNEEIKVQFAKPF 215		
Db	206	VGPVLEEMQR-----TRTN-NLDERVLTPIQLDIIICRQSLT-CPLLN-----SANF 251		
Qy	216	VYDNNSILRVSPESHDPNRFQSLSEVAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFY 275		
Db	252	YTWKSSCYRIPTA-----AGQALDLEGKEMWTGFFSSAHIASNYRPLLNIDVAHTAFY 305		
Qy	276	NAPKMSLLDYLILLIVD-----PQ 293		
Db	306	KT-RITVLQFQCDVLNERTSKPNENNPRGPGPGGPGYGRGGRGGSGYGNFGNRP 364		
Qy	294	SCN-----DDVRKD-----LKTCLMAGK---MTIRQAAARPIRQLLENLKL 331		
Db	365	GANVRDDFGGNGLTFTMDTLSRDTQLSSFETRIFGDALIRGMKIRAAHRP----- 413		
Qy	332	KCAEFWNDEMRLRTERHLTFLDLCEENS LVYVTKGSDRGNAKKYDITTLFKIYEENKPF 391		
Db	414	NAIRVYKVNLSQLPADKLMFQIGIDEEGRQV---VGSVAD-----YFSEKYGP 457		
Qy	392	IEPFLPLAVKVKSGAKAYAVPMEHLEVHEKPORVKNRIDLVMQDKFLKBATRKPHDYKEN 451		
Db	458	LKYPLCLVHVPPTNRINFLPWEHCLLT-DSPQKNKKGMSEKQTSAAIIKAAAVADATQREDR 516		
Qy	452	TLKMLKEIDFSBSBLNFVERFGLCSKLQIBCPQKVLKEPMLVNSVNEQIKMTPIVRIGFQ 511		
Db	517	IKQLAAQASFGTDP--FLKBFGVAVSQMIQTARVIOPP-----PIMFGGN 561		
Qy	512	EQLNVVPEKELCCAVF-----VNETAGNP-CLEENDV---KFYTELIGGCK-- 556		
Db	562	NRSVNPV-----VFPKQGSWTMDNQTLMPATCERSYSMALVDPDQTSLTQTFQCSL 613		
Qy	557	-FRGIRIGANENRGAQSIWYDATKNEVAFYKNCPLNTGIG-RPEIAATEAKNMPERLPDK 614		
Db	614	TMKATAMGNPWRPMDLVKYGRSKBDV-----CLTFTEIADERYVTMTVCD----- 659		
Qy	615	EQVLMTIITIIISKQLNAYGFVKHYCDHTIGVANGHITSETVTKALASLRHEKSGKRIFYQ 674		
Db	660	-----CIIIVLOSKNSDIYMTVKEQSODIVHGIMSCVLMKNVSRP-----TPATCAN 706		
Qy	675	IALKINAKLGGINQELDOWSEIAETISPEEKERRKTMPLTMVYGVIGDVTHPT-SYSGIDY-SI 732		
Db	707	IVLXLNKKMGKINSRI-----VADKIITNKYLVDPQTVVYGVIGDVTHPTQAEEMGNMPSV 759		
Qy	733	AAVVASINPGGTIYRNMLVTQEECCPERAVAHGERTDILEAKFKVLLEPFAENNDRA 792		
Db	760	AAIIVANDLIPQSGANVYKQKKCR-----ESVYV---LLDAIRERIITFYR-----HTKQK 808		
Qy	793	PAHIVWYRDGVSDEMLRVSHDELRSLSKSEVKQPMSEDRGEDPCKPTEFVIOKRNTRL 852		
Db	809	PARIIVTRDGVSEQGFSEVLREEIQSIRTACLAI-----AEDFRPPTIYIVQKRHHARI 863		
Qy	853	LRRMEKDKPVVNKDLTTPAETDVA VAAVKQWEEDMKESKETGI VNPSSPQTTVDKLLIYSKYK 912		
Db	864	PKCYQNDMWGKAKNVP-----GTTVDGTGIVSPEG 893		
Qy	913	PDPFLASHGHVGLTSHRPGHTVMTYDDKMGODEVYKVTYGLAFIUSARCRKPISLPVPVHY 972		
Db	894	PDFYLCSHYGQGTSHRPARYHVLIDDECKFTADEIQSITYGCHTYGCTRSVSIPTPVY 953		

[illegible]

Db 609 ---FVDNLIDRCWRLGWMEA-----PIVYKTSRMB-----TSLNGNAIEEL---- 647

QY 604 AKMFERLPDKE---QKVLMPFIISKRQLNAGFVKHYCDHIGVANQHITSETVTYKALA 660

Db 648 LRSVIDEASRKGGARPTLVLCAMSRKD-DGYKTLKWAETKGLVTCFLTPAT---- 702

QY 661 SLRHEKSGRIIFVOIALKINAKLGINQELDSEIAEISPEEKERRKTMPLTWYVGIDVT 720

Db 703 -----KGGQYRANLALONNAKVGGSNVEL---MDTFSFFKKEDE-----VMTFGADVN 748

QY 721 HPTSYSGIDYSIAAVASIN-PGGTIYRNMIVTQEECRPGERAVAHGRETDILEAKFYK 779

Db 749 HPAARDKMSPSIVAVVGTLMWPEANRYAARVIAQ-----PHRKEEIQFGDACLE 798

QY 780 LLREFAEENDNAPAHIVYRQGVSDSEMLRVSHDELRLSKSEVKQFMSERGDEGPBK 839

Db 799 LVKAHVQATGKR-PNKIVIFRQGVSDAQDFMLNVLELLDVK-----LTPEKNGYN--PKI 850

QY 840 TFIIVQKRHTLLRRMEKDKPVNKKDLTPAETDVAVAAVKQWEDMKESKETGIWNPSS 899

Db 851 TVIVAQRHQRTRFF-----PATNND-----GSDKG--NVPS 879

QY 900 GTTVDKLIYSKYKDFPFLASHHGVLTSPGHYTYMYDDKMGSDQEVYKMTYGLAFLSAR 959

Db 880 GTVWDTKVHPYEDFYLCSHHGIGTSKPTHYTLWDELGFTSDQVQKLIFEMCFTR 939

QY 960 CRKPTSLPVVHYAHLSCAKELRYTYKEHIGDYAQR 999

Db 940 CTKPVSLVPVYADWAFRGRMYHEASSRE--KNFKQPR 977

RESULT 11

Ts2134

Zwille protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004

C:Accession: TS2134

R:Laux, T.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z25975

A:Accession: TS2134

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-988 <LAU>

A:Cross-references: UNIPROT:Q9XGWI; UNIPARC:UPI000016DA03; EMBL:AJ223508; PIDN:CAAL14239.

C:Genetics:

A:Gene: zwille

A:Map position: 5

C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 9.1%; Score 488; DB 2; Length 988;

Best Local Similarity 22.7%; Pred. No. 1.5e-19;

Matches 221; Conservative 155; Mismatches 377; Indels 222; Gaps 38;

QY 49 FSSKYVDREYVEYEVQMTKEVLNRKPGKPPKTEIPIDRAKLFQHLRHEKKOTDFIL 108

Db 148 FIADIPTKDLQNDYDVTITEVSSKSVNRAI-----IAELVRLY-----KESDLGR 192

QY 109 EDYVFDEKDTVYVSVCLNTVTSKMLVSEKVKKD-----SEKDEKDKLEKKILYTMILTYR 164

Db 193 RLPAVDGRKSLVTAGELPPTWKEF--SVKIVDEDDGIIINGPKRERSYKVAIKF---VARA 247

QY 165 KKFHLN--PSRENPEKDEANRSYKFLKNVMTQKRYAP-----FVNEEIKVQFKNFYVD 218

Db 248 NMHILGEFLAGKRCADPQBAVQILDIVLRELSVK-RFCVPGRSFFSPDIKT----- 297

QY 219 NNSILRVPSFHDNPRFOSLEVAPRIEAWFGIYIGIKELFDGEPVLNPAIVDKLFYN-A 277

Db 298 -----PQRLGSL-----ESWCQFYQIRPTQWGLS-LNIDWASAFAIEPL 337

QY 278 PKMSLLDYLL---LIVDPQSCNDVVRKDLKTKLMAGKMTIROAARPRIRQLLENKLKKA 334

Db 338 FVIEFVAQLLGDVLSKPLSDSRVK--IKKGLRGVKVVEVTHRANVRRKYRVAGLTTQ-- 393

QY 335 EVDNEMSLTBRHLTFLDLCSENSIVYKVTGKSDRGRNAKKYDTTLFKIYBENKKF-IE 393

Db 394 -----PTRELMP--PVDEN-----CTWKS-----VIEFYQEMYGFTIQ 424

QY 394 FPHLPLVYKSGAKGYAVPMHEHLEVHEKPYQRYNRIDLYMODKFLKRAATKPHDYKENTL 453

Db 425 HTHLPCLQVGNQKASYLPMPEACKIVE-QQRYTKRLNEKQITALLKVTQCORAEQGRNDIL 483

QY 454 KMLKELDFSSSELNFERPGLCSKLQMIETCPQVKLKEPMLVNSVNEQIK-MTPVI---RG 509

Db 484 RTVQHNAYDQDP--YAKEFGMNISEKLASVEARILPAPWLKYHENGKEDCLPQVQQNM 541

QY 510 FOEKQLNVVPEKELCCAVFVNNETAGNCPCLBENDVVVKFYTELIGGCKFKGIRIGANENRG 569

Db 542 MNKQWINGMTVSRWACVNF-----SRVQENVARGFQCNELQOMCEVSEMEPNPFPVP 594

QY 570 AQOSIMYDATKNEYAFYKNCNLNTGIGRFEIAATEAKNMFERLPDKQKQVLMFIIISKRQL 639

Db 595 IYSARPQDQVEKALKHVYHTSMNKTGK-----ELELLALLPDNNGSL----- 637

QY 630 NAYGFVKHYCDHIGVANQHITSETVTYKALASLRHEKSKRIFYQIALKINAKLGINQE 689

Db 638 --YGDLKRICETELGLISQCCLTKEHFKI-----SKQYLADVSLKINVMQGRNTV 686

QY 690 LDMSIAETISPEEKERRKTMPL-----TWYVGIDVTHPTSYSGIDYSIAAVASIN-PGG 743

Db 687 L-----VDALSCR-----IPLVSDIPIIFQADVTHPENGBRESSPSIAAVASQDWPEV 735

QY 744 TIYRNMIVTQEECRPGERAVAHGRE-----RTDILEAKFVKLLREFAEENNDN 790

Db 736 TKYAGLVCAQ-----AHRQELIQDLYKTQDVPVRGTVSGGMIRDLLISFRKATGQ 785

QY 791 RAPAHIVYRQGVSDSEMLRVSHDELRLSKSEVKQFMSERGDEGPBKY----TFIVIOK 846

Db 786 K-PLRIIFYRQGVSEGFYQLLYELDAIRKACASL-----EPNPQPPVTFIVVQK 835

QY 847 RHNTLLRRMEKDKPVNKKDLTPAETDVAVAAVKQWEDMKESKETGIWNPSSGTTVDKL 906

Db 836 RHHTLRFPANNHRDK-----NSTDRSGNILP--GTVVDYK 867

QY 907 IVSKYKDFPFLASHHGVLTSPGHYTYMYDDKMGSDQEVYKMTYGLAFLSARCRKPISL 966

Db 868 ICHPTEPDFYLCSHAGIQGTSRPAHYHVLWDENNNTADGIQSLTNNLCVYARCTRSVSI 927

QY 967 PVPVHYAHLSCAK 981

Db 928 VPPAYYAHLAARFR 942

RESULT 12

Ts1568

argonaute-like protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: Ts1568

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z21991

A:Accession: Ts1568

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-834 <WOO>

A:Cross-references: UNIPROT:O74957; UNIPARC:UPI000012D2A7; EMBL:AL023705; PIDN:CAAL19275

A:Experimental source: strain 972h-; cosmid c736

C:Genetics:

A:Gene: SPDB:SPCC736.11

A:Map position: 3

A:Introns: 43/3

C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 9.0%; Score 484; DB 2; Length 834;

Best Local Similarity 21.8%; Pred. No. 2e-19;

	Matches	212; Conservative	165; Mismatches	372; Indels	222; Gaps	39;
Qy	38	KKVLLVNMFKPSKIVDRYYEYVKMTKEVLNR-----KPGKPPFKCTEIPIDRAKL	92			
Db	20	KQITLKANFFQI-----ISLPNETINQYHIVIGDGRVPRK-----QSQL	59			
Qy	93	FWOHLRHEKKQTDFILEDYVDEKDTVYVSVCLNTVTSMKLVSEKVKVCKSEKKDKOLE	152			
Db	60	IW-NSKEVKQYFGSSWNVSVDGRSMCWSKGDIADGTIKVNGISESHPRETEFISQKS-S	117			
Qy	153	KKILYTMILTYRKXPHLNFSGRENPEKDEEANRSYKFLKNVNMVTKRYAPFVNBEIKVQFA	212			
Db	118	KINLHTLSQFVNSKY-----SSDPQLSSIMFLDLLLKKK-----PSETLFGFM	161			
Qy	213	KNFYVDNNSILRVPDESHPDNRPQOSILEVAPRIEAWFGIYIGIKELFDGPVPLNFAVDK	272			
Db	162	HSPTGTENGW-----SLGGGVBAWKFYQSIIRP-NQGFMVSNVDISSL	203			
Qy	273	LFYNAPKWSLLDYLLLLIVDPQSCNDVDYRKDKTKLMAGKMTIRQAAAPRIROLLENLKLK	333			
Db	204	AFWR--NDSLLOILMEYTDCSNVDDLTRFDLK-----RLSRKPRFLKVT	245			
Qy	333	CAEVDNEMSGRLTERHITFTLDLCENSLUVKYVTKSD-----RGRNAKYDVTTLFK	383			
Db	246	C-----QHRNNVGTDLNR--VYSIEGFSSKASDSPFVRLNGEEQKISVAE	291			
Qy	384	IYENKKP-IETPHLLPVKYSKAKEYAVPMHELVHEKQRYKNRIDLVMQDKFLKRAI	442			
Db	292	YFLENHNVRLOYPNLPCLIKVNGA--MLPIEFQCV-VKQORYAKLNSDQTANMIRFAV	347			
Qy	443	RKPHDYKENTLKMKLKELDFSEELNPFVERFGLCSKLQMIQCPGVKLPKPMPL--VNSVNE	499			
Db	348	QRPERVQOIDDVFHQMDWTDTP--YLTQYGMKIQKQMLEVPARVLETPSIRYGGDCIER	405			
Qy	500	QIKMTPIVRGQEKQLNVPEKEL-CCAFVFNVTAGNCPGLENDVVVKFYTELIGCGKFR	558			
Db	406	PVSGRWNLRG--KRFLDPPRAPITRSWAVMCFITSTRRLPMRGIEFNLTQYVQTL-----	456			
Qy	559	GIRIGANENGAQSIMYDATKNEYAFYKNCITLNTIGIRFSAIAT---EAKNMPERLPDKE	615			
Db	457	TSLGINFVMMKPPVLADIIRGSVE-----ELCITLYKKASQOVGNAPD--	499			
Qy	616	QKVLMTIISKRLQNLAYGFVKHYCDHTIGVANGHITSETYTKALASLRHEKSGKRIFYQI	675			
Db	500	--YLPFLDKNSPEPYGSIKRVNTMLGVPSCAISKHILQS-----KPQYCANL	547			
Qy	676	ALKINAKLGGINQELDWESETABISPEBERKKTWPL---TMVVGIDVTHP-TSYSGIDY	730			
Db	548	GMKINVKVGGIN-----CSLIP-----KSNPLGNVPTLILGGDVYHPGVGATGV--	591			
Qy	731	SIAAVVASINPGGTIYRNMIIVTQBECPRGRAVAHGRTDILEAKPVCLLRBEAENNND	790			
Db	592	SIASIVASVDLNGCKYTAVERSQ---PRHQEVIEGAK--DIV---VYLQGFRAWT-K	640			
Qy	791	RAPAHIVVYVRGVSDEMLRVSHDELSLKSEVKQFMFSRDEGDPEPKYTFYVIQKRHT	850			
Db	641	QOPQRIIYFRDGTSGEQFLSVINDELSQIKEACHSLSPKN-----PKILVCTTQKHHH	695			
Qy	851	RLLRMEKDPVNVKDLTPAETDVAAVAKQWBEEDMKESKETGIWNPSSGTTVDKLVSK	910			
Db	696	RFP-----IRNK-----SDGRNG--NPLPGTITIEKHVTHP	724			
Qy	911	YKDFDFLASHHGVLGTSRPGHYTMVDDKGMGSODEVYKMTYGLAFISARCKRPISLPVV	970			
Db	725	YQYDFYLLSHPSLQGSVPVHYTVLHDEIQWPPDQFTCLCNLCYVVARATSAVSLVPPV	784			
Qy	971	HYAHLSCAK	981			
Db	785	YYAHLVSNLAR	795			

RESULT 13
T25164

hypothetical protein T23D8.7 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T25164
R/Wild, A.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19989
A/Accession: T25164
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-898 <WTL>
A/Cross-references: UNIPROT:Q9XV13; UNIPARC:UPI000007743E; EMBL:Z81128; PIDDN:CA
A/Experimental source: clone T23D8
C/Genetics:
A/Gene: CBSP: T23D8.7
A/Map position: 1
A/Introns: 40/3; 63/2; 188/2; 508/3; 552/2; 785/3; 849/3
C/Superfamily: rabbit translation initiation factor eIF-2C

Query Match	8.8%	Score 473	DB 2	Length 898
Best Local Similarity	22.7%	Pred. No. 9.3e-19		
Matches	242	Conservative 160	Mismatches 366	Indels 296
Gaps	46			
Qy	2	SSNPPELEKGF-----YRHSLDPE---MKWLARPPGKCDGKEYEKKVLLVNV	45	
Db	31	SINVPSELENFLSSSGSRVSDDLYLTHPIEENREPPKLGKPLPSTTGRF---LSLGLAN	86	
Qy	46	WKFQF--SSKIYDREYTEYEVKMTKEVLRNPKGPPFKTEIIPDRAKLFWQHLRHEKKQ	103	
Db	87	HFOITNGSIIHQYIRFPDDPIESKKLNR-----TLTTLTQLEQ	124	
Qy	104	TDFILE-DYVDFEKKDTVYSVCLNTV---TSKMLVSEKVVVKDSEKDKDLKKIILYTM	159	
Db	125	NPGLIECPPLVFDGIHTVYSTELINVKEVNNSVINAGVNVTKESP-----LFLK	174	
Qy	160	ILTYRKXPHLNFSGRENPEDKEBANRSYKFLKNVMTOKRVYAPFVNEBIKQVPAKNFVDN	219	
Db	175	YLTHVDSFLDTHIITGNODQ-----NOKLRMMHAIDTVFRQTSIGNF-----	217	
Qy	220	NSILRVESPESHDPNRRFQS-----LEVAPRIEAMFGIYIGIKELPDGPEVPLNFAIVDK	272	
Db	218	HAVLOQSFSSIAQNSAIEPESHGLGWGTNLGVREVCYGFYQNVVETFD-TLTWNLDVATT	276	
Qy	273	LFYNAPKMSLLDY-----LLLIIVDPOSCNDVVRKDLTKLMAGRWTTIRQAARPIRQ	324	
Db	277	TFYR--PVALVEFLAEILEVPLATVTDGRSLSDVQKKFNREVAGLVKVFETHCSPRFR	334	
Qy	325	L-----LENLKLCAEVDNEMSRLETHLTFDLCEENSLVYKVTGKSDRGRNAKY	377	
Db	335	VARCTWKPTENISFHLSETAGNQDSP-----LSLVE-----YYK-----RRY	372	
Qy	378	DTTLFKIYBENKFIFFPHLPVVKVSGAKYAVPMEHLEVHEHKPORYKNRIDLVMDQKP	437	
Db	373	NIDL-----TYKHLPCIEV-GRTRCFLTLELCYV-VSGORCTIKKLN-EQOIAN	418	
Qy	438	LKRAIRKPHYKENTLKLKELDFFSSEELNFVERFGLCSKLOMIECPGKVLKPEMLV---	494	
Db	419	LIPATSGNATERQNAVMSLQNLKMDNDVNAV-KFGLKVEAQLLKTEGRVLPVPRLLYRS	477	
Qy	495	-NSVNEQIKWTP-----VIRFQEQKLNVV-----PE--KELCCCAVFVNV-----	531	
Db	478	PNLKRQECVTPNNGTMDMRGNFYSGIQIREWAI VCPASPEIIGEASMSFVRNLVNVA	537	
Qy	532	ETAGNFCLENDVVKYFTELIGCKEPRGIRIGANENRGAQSIYMDATKNEYAIFYKNCTLN	591	
Db	538	SEIGMPFLBEH---RF-----CRY-----	553	
Qy	592	TGIGRFEIAATAKNMFERLPDKEQKVLMIIFIIISKQLNAYGFVKHYCDHTIGVANOHIT	651	
Db	554	-----APDQTVKLEHLNEQYNIQLVLCIVPGKSV-VYGELEKRRK-GELLGTTTCQVR	604	
Qy	652	SETVTYKALASLRHEKSGKRIFYQIALKINAKGLGQINQELDWSBIABISPEEKERRITMPL	711	

Db 605 SONVSKA-----SPHTLSNLQWINKSLGGINVL-----SPQSLNBPV-- 646
Qy 712 TMYGIDVTH-----PTSYSGIDYSIAAVASINPGGTYIRNMIVTQECRPGERA 762
Db 647 -LFICHLTRSLASSDSTSIACHDSSIACLVSGMDGHPQFQPIFRTPR----- 698
Qy 763 VAHGRERTDILAKFVKLLRPAEN---NDNRPAHIVVYRQVSDSEMLRSHVSHDLRSL 819
Db 699 --HQTIVDMCE-----MTREALINFRKSTGPKPHKIIIRAGIADVTVDIEMQELRAV 751
Qy 820 KSEVQFMSRDEGDEPEPKYPIVIOKRNTHLLRMKDKPVPVNDLTPASTDVAAVAV 879
Db 752 RDACA--MIEYGF--QPCITFGLDVTHTLFAANEKOR--VGNSONVPAGTLV----- 800
Qy 880 KQWEDMKESKETGIVNPSGGTVDKLVSKYKFPDFFLASHHGVLTGSRPGHYTMYDDK 939
Db 801 -----ETGI-----TVNVL-----FEFYLVSAGIOGTSRPTKYVVMWDDN 836
Qy 940 GMSODEVYKMTYGLAFLSARCKPISLPPVHYAHLSCAKEL 983
Db 837 SIPSADIHEMTYQLCHTQSRCTRSVSPSPVYAKLVQAQAKIL 880

RESULT 14
A:Argonaute (AGO1)-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
R:Accession: A84668
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Status: preliminary
A:Accession: A84668
A:Molecule type: DNA
A:Gene: At2g27040
A:Map position: 2
A:Cross-references: UNIPROT:Q9ZVD5; UNIPARC:UPI0000179101; GB:AE002093; NID:g3885334; PI
C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 8.0%; Score 428.5; DB 2; Length 930;
Best Local Similarity 22.0%; Pred. No. 3.le-16;
Matches 218; Conservative 167; Mismatches 381; Indels 223; Gaps 40;
Qy 38 KKVLLVNMVFPSSKIYDREYVEYEVKMTKEVLNRKPGKPPPK-TEIPIDRAKLFQW 96
Db 67 QKIPLLTHFKVDVANLQGHFFHYVALFYD-----DGRPVEQKGVRKILDKV----- 115
Qy 97 LRHEKKQTDFILEDYVFDEKOTVYS-----VCLRLNTVTSKMLYSE----- 136
Db 116 --HQTYSLDLDCGEKAYDGEKLTFTYGALPSNKMDFSVLEBVSATSKDFVSBRANGNSP 173
Qy 137 --KVYKSEKKDEKOLEKILYTMILYTKKFLHFNFSRENPEKDEEARSFKLKNVMT 194
Db 174 NGNESPSDGRKRLRPNRSKNFRVSIYAAKIPLQ-ALANAMRGESNSQEAIR----- 228
Qy 195 OKRVYAPFVNEBIKQVAKNFVYDNNLSILRVPESE--HDPNRFQESLEVAPRIEAWFGIY 252
Db 229 -----VLDIILQHAH-----RQGLLVRSQSFHNDPTNCE---PVGGNLSGRGFH 272
Qy 253 IGKELFGPEVFNFAIDKLFYNAPKMSLLDYLIVDPOSCNDVDVRKDLTKLMAGKM 312
Db 273 SSFRIT-QQGMSLNDVTTMIKFG--PVVDFLI-----ANQNARDPYSIDWSKAKR 322
Qy 313 TIROQAPRIQOLLNKLKCAEVDN-EMSLRTHLFTLDCENSLVYKVKTSRGR 371
Db 323 T-----LKNLRVKVSPSQGEFKITGLSDKP-----CREQTFELKRNPNENG 364

Qy 372 RNAKYDTTLFKI--YBENKKPIEPHP---LPLVKVSKGAKAYAVPMHELVHEKQRYK 426
Db 365 ----EFETTEVTADYFRDTHIDLOYADLPICINVGPKRPTYIPLE-LCALVPLQRYT 419
Qy 427 NRIDLVMQDKFLKATRKPHDYKENTLKWLEDFSSBELNVPFRGLCSKLQMIKCPQK 486
Db 420 KALTTFQRSALVERSKRQPKQBRMTVLSKALKVSNYDABPL--LRSCGISISSNFTQVEGR 477
Qy 487 VLKEMLVNSVNEQIKMTFVIRGFOEKQLNVVPEKELCCAVFVNNETAG-NFCLRENDVV 545
Db 478 VLPAPKLKMGCGSETFPFRNGRWNFNKKEF-VEPTK--IQRVVVVNFPSARCNVRQVVDLLI 534
Qy 546 KFYTELIGCGKFRGIRIGA-----NENRGAQSIMYDATKNEYAFYKKNCTLNTGIGRF 597
Db 535 K-----IGSK--GLEIASPPQVFPFEGNQFRAPPWI----- 564
Qy 598 EIAATEAKMFERLPDKQKVLMIIFII---SKRQLNAYGFVGHYCDHTGTIGVANQHITSET 654
Db 565 ----RVENMPKDIQSKLPGVPQFILTCLVLPDKKNSDLYGPMKKQLNTEFGI----- 610
Qy 655 VTKALASLRHEKSKRIFYQIALKINAKLGGINQELDNSEIAEISPEEKERRKTMPLTWY 714
Db 611 VTQCNAPTR--OPNDQYLTNLLKINAKLGGINSMLSVERTPAFTVISK-----VP-TII 662
Qy 715 VGIDVTHTSYSGIDYSIAAVASIN--PGGTYIRNMIVTQECRPGERA--AHGR 767
Db 663 LGMDVSHGSPQSDVPSIAAVSSREWPLISKYRASVRTQ-----PSKAEMIESLVKNGCT 718
Qy 768 ERTDILEAKFVKLLRFEANNDNRPAHIVVYRQVSDSEMLRSHVSHDELRSUKSEVKQPM 827
Db 719 EDDGIK---ELLVDFVTSNNKRPKEHIIIPRDGVSESQFNQVLNIELDQIIEACKLLD 774
Qy 828 SERGEDPEPKYTFVIOKRNTHLLRMKDKPVPVNDLTPAETDVAVAVKQWEDMK 887
Db 775 ANWN-----PKPILLVAQKNHHTKFF-----QPTSPENVP----- 805
Qy 888 ESKETGIVNPSGGTVDKLVSKYKFPDFFLASHHGVLTGSRPGHYTMYDDKMSQDEYV 947
Db 806 -----GTIIDNKICHKPNDFYLCAHAGMIGTTRPTHYHVLVDEIGFSADELU 853
Qy 948 KMTYGLAFLSARCKPISLPPVHYAHL 976
Db 854 ELVHSLSVYQRTSSTSAISVAPICVAHLA 882

RESULT 15

T01113
translation initiation factor eIF-2C homolog T21L14.12 - Arabidopsis thaliana
N:Alternate names: Argonaute (AGO1)-like protein
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01113; D84739
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Cressby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,
submitted to the EMBL Data Library, December 1997
A:Description: Arabidopsis thaliana chromosome II BAC T21L14 genomic sequence.
A:Reference number: Z14209
A:Accession: T01113
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-887 <ROU>
A:Cross-references: UNIPROT:Q48771; UNIPARC:UPI00000AB808; EMBL:AC003033; NID:g2702261;
R:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Status: preliminary
A:Accession: D84739
A:Molecule type: DNA
A:Residues: 1-887 <STO>
A:Cross-references: UNIPARC:UPI00000AB808; GB:AE002093; NID:g2702284; PIDN:AAB91987.1; <

C:\Genetics:

A:Gene: A02932940; T21L14.12

A:Map position: 2

A:Introns: 61/3; 130/2; 203/2; 248/1; 281/3; 314/2; 367/3; 406/3; 442/3; 463/3; 501/3; 5

C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 7.58; Score 402; DB 2; Length 887;

Best Local Similarity 22.48; Pred. No. 8.8e-15;

Matches 227; Conservative 160; Mismatches 358; Indels 268; Gaps 45;

```
QY 42 LLVNWKFSSKIYDREYIEYEVKRLKPKPPKTEIPDPRAKLFWQHLRHEK 101
Db 41 LCTNHFNVSVRQPDVVVFYQYTVSITTEGDAVDGTGSRKL-----MDQLP-----KT 88

QY 102 KQTFDILEDVDFEKTQTVSVCRLL--NTVTSKMLVSEKVVKDKSEKDE-----KDL 152
Db 89 YSSDLGKRLAYDGEKTLTVGPLQNEFDLVIVGFSKRGVSDGSSGTCRKS 148

QY 153 KKILYTWILTYRKKFHLNFSRENPEKDEEANSYKFLKNVM-TQKVRYAPFVNEEIKVQ 211
Db 149 RSLPR---SYKQIIR--YAAEIP-----LKTVLGTORGAYTPDKSAQ----- 186

QY 212 AKNFVYDNNLSILVPSFSDPNRFEQSLEVPAPRIEAWFGIYIGIKELF---DGE 263
Db 187 -----DALRV-----LDIVLQQAARCGCLLVROAFPHSDGHPMKVG 224

QY 264 -VLNFAIVDKLP---YNAPKMSLLDYLLLVDPQSCNDDVRKDLTKLMAGKMTIRQAAR 319
Db 225 GVIGIRGLHSSFRPTGGSLNLDVSTTMILEPGVPVIEFLKANQSVETPRQIDWIKVA 284

QY 320 PRIQRLLENLKLKCAEWDNEMSLRTERHTFLDL-----CRENSLVYKVTGKSDRGRN 373
Db 285 -----MLKMRVKA-----THRNMEFKIIGLSSKPCNQQLFSMKI---KDG 325

QY 374 AKKYDTTLFKIYBEN-KKFTFPHPLVVKYGAKEVAPME-----HLEVHEKPKQYKN 427
Db 326 VPIREITVDYDFYKQYTEPISSAYFPCLDVGKPDPRNLYLPLEFCNLVSLQRYTKPLSGRQ 385

QY 428 RIDLVNQDKPLKRAKTRKPHDYKENTLKMKLKD-----FSSEELNFERFGLCSKLQMI 482
Db 386 RVLLV-----ESSRQKP-----LERIKTLANDAMHTYCYDKDPFLAGGGSIEKEMTQ 432

QY 483 CPCKVLKEPMLVNSVNEQIKMTFVIRGQEQKQLNVVPEKELCCAVFVNETAGNCPLEEN 542
Db 433 VEGRVLPKPMLKFGKNEDFQPCNGRWNFNKML---LEPRAIKSWAIVNFSF--PC--- 483

QY 543 DVVKFYTELIGGCKFGIRIGANENRGAQSIMYDATKNEYAFYKNCNTLNTGIGRPEIAAT 602
Db 484 DSHISRELI-SC---GMRKGIEDRPFALVEEDPQ-----YKKA-----GPVERVEK 527

QY 603 EAKNMPERLPDKEQKVLPMFIILSKROLNAYGFVKHYCDHTIGVANQHIITSETV--TKALAS 661
Db 528 MIATMKLKPDPDPHFILC-ILPERKTSDIYGPWKIC-----LTEEIGHTQCICP 576

QY 662 LRHEKSKRIFYQIALKINAKLGINQELDWSIAISPEEKERRKTMPL-----TWYVG 716
Db 577 I---KISDQVLTNVLKINSKLGINSLLG-----IEYSNIPLINKIPTLIILG 622

QY 717 IDVTH-PTSYSGIDYSTAAVVAS-INPGGTIYRNMIWVTO-----EECRPGE 760
Db 623 MDVSHGPPGADVP-SVAAVGSKCWPLISRYRAAVRTQSPLEIMDSLPQPIENTEKGD 681

QY 761 RAVAHGRETDILEAKFVKLLREFAEENNDRAPAHIVVYRDGVSDEMLRVSHDELRS 820
Db 682 NGIMN-----ELFVEFYRTSRARKPKQIIIFRDGVSESQEQV-----LK 721

QY 821 SEVKQFMS---ERDGEDPEPKYTFIVIOKRNTLLRRMEKDKFPVVKNDLTPAETDVAVAA 878
Db 722 IEVDQIIKAYQRLGESDVPKFTVIVAQKNHTKLF----- 756

QY 879 VKQWBEEDMKESKETGVNPSSTGTTVDKLIIVSKYKFDFFFLASHHGVLGTSRPGHYTVMYDD 938
Db 757 -----QAKGPENVPAGTVVDTKIVHPNTYDFYMCNHAGKIGTSRPAHYHVLDE 805
```

QY 939 KGMSQDEVYKMTYGLAF-----LSARCRKPISLPVPVHYAHLSCSEKAKE 982
Db 806 IGSPDDQLQNLHLSYKLIINSIFNVSSLLCVFVLSV-APRVYAHLAQAQVAQ 857

Search completed: July 5, 2006, 13:34:44
Job time : 56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM protein - protein search, using sw model

Run on: July 5, 2006, 13:25:16 ; Search time 315 Seconds
(without alignments)
2995.289 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 MGSNPFLEKGFYRHSLEDP.....RHEMFPLQTNVYKPGMSFA 1020

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5349	100.0	1020	Q9XU82 CAEEL	Q9XU82 caenorhabdi
2	3241	60.6	976	Q60V50 CAEER	Q60V50 caenorhabdi
3	604.5	11.3	997	AGOL ARATH	Q9sjk3 arabidopsis
4	577	10.8	857	I2C1 HUMAN	Q9ul18 homo sapien
5	573	10.7	863	Q4RKH3 TETNG	Q4rkh3 tetraodon n
6	571.5	10.7	848	Q8CGU0 MOUSE	Q8cgu0 mus musculu
7	571	10.7	857	I2C1 MOUSE	Q8c9j1 mus musculu
8	568.5	10.6	1055	Q7Y001 ORYSA	Q7y001 oryza sativ
9	567.5	10.6	891	Q7PZ41 ANOGA	Q7pz41 anopheles g
10	565.5	10.6	862	Q6DCX2 XENLA	Q6dcx2 xenopus lae
11	565	10.6	840	I2C2 RABIT	Q77503 oryctolagus
12	565	10.6	860	I2C2 MOUSE	Q8c9j0 mus musculu
13	565	10.6	871	Q6DUB9 XENLR	Q6dj59 xenopus tro
14	564	10.5	860	I2C2 RAT	Q9qz81 rattus norv
15	564	10.5	883	Q4VAB3 MOUSE	Q4vab3 mus musculu
16	563	10.5	860	Q6QME8 BOVIN	Q6qme8 bos taurus
17	563	10.5	1024	Q61PV1 CAEER	Q6lpv1 caenorhabdi
18	560.5	10.5	950	Q9V312 DROME	Q9v312 drosophila
19	560.5	10.5	984	Q9V6V6 DROME	Q9v6v6 drosophila
20	558	10.4	859	Q4S7L5 TETNG	Q4s7l5 tetraodon n
21	557	10.4	1002	Q20578 CAEEL	Q20578 caenorhabdi
22	557	10.4	1010	Q3LFR7 CAEEL	Q3lfr7 caenorhabdi
23	556.5	10.4	851	I2C2 HUMAN	Q9ukv8 homo sapien
24	555.5	10.4	703	Q571J6 MOUSE	Q571j6 mus musculu
25	555.5	10.4	860	I2C3 HUMAN	Q9h9g7 homo sapien
26	555.5	10.4	860	I2C3 MOUSE	Q8c9j9 mus musculu
27	555	10.4	861	I2C4 HUMAN	Q9hck5 homo sapien
28	555	10.4	861	Q5VXF0 HUMAN	Q5vxf0 homo sapien
29	554.5	10.4	782	Q5TA58 HUMAN	Q5ta58 homo sapien
30	553	10.3	861	Q4VBD7 MOUSE	Q4vbd7 mus musculu
31	553	10.3	951	Q8BTF4 MOUSE	Q8btf4 mus musculu

32	552.5	10.3	860	2	Q5ZLG4 CHICK	Q5zlg4 gallus gall
33	550.5	10.3	860	2	Q3TBF7 MOUSE	Q3tbp7 m nod-deriv
34	547.5	10.2	1038	2	Q5Z5B2 ORYSA	Q5z5b2 oryza sativ
35	546.5	10.2	1032	2	Q7JLZ2 CAEEL	Q7jlz2 caenorhabdi
36	546.5	10.2	1035	2	Q21079 CAEEL	Q21079 caenorhabdi
37	544	10.2	861	1	I2C4 MOUSE	Q8c9f8 mus musculu
38	542.5	10.1	945	2	Q4SVB6 TETNG	Q4sve6 tetraodon n
39	542	10.1	884	2	Q4KL6 XENLA	Q4klv6 xenopus lae
40	540.5	10.1	891	2	Q8EB39 CAEEL	Q8eb39 caenorhabdi
41	540.5	10.1	910	2	O16720 CAEEL	O16720 caenorhabdi
42	539	10.1	794	2	Q5ZMW0 CHICK	Q5zwm0 gallus gall
43	537	10.0	1040	1	Y043 CHICK	P34681 caenorhabdi
44	536.5	10.0	1034	2	Q7XTS4 ORYSA	Q7xts4 oryza sativ
45	536.5	10.0	1058	2	Q851R2 ORYSA	Q851r2 oryza sativ

ALIGNMENTS

RESULT 1
Q9XU82 CAEEL
ID Q9XU82 CAEEL PRELIMINARY; PRT; 1020 AA.
AC Q9XU82; Q9U6Q1;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2001, sequence version 2.
DT 07-FEB-2006, entry version 26.
DE Hypothetical protein rde-1 (RNA interference promoting factor RDE-1).
GN Name=rde-1; ORFNames=K08H10.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20004389; PubMed=10535731; DOI=10.1016/S0092-8674(00)81644-X;
RA Tabara H., Sarkissian M., Kelly W.G., Fleenor J., Grishok A.,
RA Timmons L., Fire A., Mello C.C.;
RT "The rde-1 gene, RNA interference, and transposon silencing in C.
elegans."
RL Cell 99:123-132(1999).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; Z83113; CAB05546.2; -; Genomic_DNA.
CC EMBL; AF180730; AAF06159.1; -; mRNA.
CC PIR; T23510; T23510.
CC Ensembl; K08H10.7; Caenorhabditis elegans.
CC WormBase; WBGene0004323; rde-1.
CC WormPep; K08H10.7; CE28243.
CC InterPro; IPR003100; PAZ.
CC InterPro; IPR003185; Pwi.
CC Pfam; PF02170; PAZ; 1.
CC Pfam; PF02171; Pwi; 1.
CC PROSITE; P850822; Pwi; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1020 AA, 118804 MW, 8E2F1A2EFC43A670 CRC64;

Query Match 100.0%; Score 5349; DB 2; Length 1020;
Best Local Similarity 100.0%; Pred. No. 1.2e-292;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSNPFLEKGFYRHSLEDPDKMQLARPTGKCDGKFKYKKVLLVNNFKSSKIYDREYYE 60
|||||
DB 1 MGSNPFLEKGFYRHSLEDPDKMQLARPTGKCDGKFKYKKVLLVNNFKSSKIYDREYYE 60
|||||

```
QY 61 YEVKMTKEVLNRKPKPKTEIPIDRAKLPWOHLRHEKKQTDIFLEDYVDFSKDTVY 120
Db 61 YEVKMTKEVLNRKPKPKTEIPIDRAKLPWOHLRHEKKQTDIFLEDYVDFSKDTVY 120
QY 121 SVCLNTVTSKMLVSEKVKVKKSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
Db 121 SVCLNTVTSKMLVSEKVKVKKSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
QY 181 EARNRYKFLKNVMTQKRVAPFVNEIKVQPAKNFYDNNSTILRVPSFHDNRPEQSLE 240
Db 181 EARNRYKFLKNVMTQKRVAPFVNEIKVQPAKNFYDNNSTILRVPSFHDNRPEQSLE 240
QY 241 VAPRIEAMFGIYIGIKELPDGEPVLNFAVDKLFYNAPKMSLDYLLIVDPSQNDVVR 300
Db 241 VAPRIEAMFGIYIGIKELPDGEPVLNFAVDKLFYNAPKMSLDYLLIVDPSQNDVVR 300
QY 301 KDLTKLMAGKMTIROAARPRIRQLLENLKLCAEVDNEMSRLETHERLTFDLDCEENSL 360
Db 301 KDLTKLMAGKMTIROAARPRIRQLLENLKLCAEVDNEMSRLETHERLTFDLDCEENSL 360
QY 361 VYKVTGSDRGNAKYDTTLTKIYEENKFTIEFPHLPLVKYKSAKEYAVPMHEHLEVE 420
Db 361 VYKVTGSDRGNAKYDTTLTKIYEENKFTIEFPHLPLVKYKSAKEYAVPMHEHLEVE 420
QY 421 KPORYKNRIDLVMDQKFLKGRATKPDHYKENTLKMKELDPSSEELNFVEREGLCSKQM 480
Db 421 KPORYKNRIDLVMDQKFLKGRATKPDHYKENTLKMKELDPSSEELNFVEREGLCSKQM 480
QY 481 IECPGKVLKEPMLVNSVNEIQIKMTPIVIRGFBQKQLNVVPEKELCCAVFVNNETAGNPCLE 540
Db 481 IECPGKVLKEPMLVNSVNEIQIKMTPIVIRGFBQKQLNVVPEKELCCAVFVNNETAGNPCLE 540
QY 541 ENDVVKFYTELIGCGKFRGIRIGANENRGAQIMYDATKNEYAFYKNCYTLNTGIRFEITA 600
Db 541 ENDVVKFYTELIGCGKFRGIRIGANENRGAQIMYDATKNEYAFYKNCYTLNTGIRFEITA 600
QY 601 ATEAKNMFRLPDKEQKQVLMFIIISKQLNAYGFVKHYCDHTIGVANOHITSETVTYKALA 660
Db 601 ATEAKNMFRLPDKEQKQVLMFIIISKQLNAYGFVKHYCDHTIGVANOHITSETVTYKALA 660
QY 661 SLRHEKSGRIFYQIALKINAKLGGINQELDWSEIAEISPEEKERRKTMPLTYVYGIDVT 720
Db 661 SLRHEKSGRIFYQIALKINAKLGGINQELDWSEIAEISPEEKERRKTMPLTYVYGIDVT 720
QY 721 HPTSYSGIDYSIAAVASINPGGTYIRNMIVTQSECRPGERAVAGHRETDILEAFVKL 780
Db 721 HPTSYSGIDYSIAAVASINPGGTYIRNMIVTQSECRPGERAVAGHRETDILEAFVKL 780
QY 781 LREFARNNDNRAPIHVVYRVDGVSDEMLRVSHDELRSLSKVKQFMSRDEGDEPEPKYT 840
Db 781 LREFARNNDNRAPIHVVYRVDGVSDEMLRVSHDELRSLSKVKQFMSRDEGDEPEPKYT 840
QY 841 FVIOQRHNTRLLRMEKDPVNVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
Db 841 FVIOQRHNTRLLRMEKDPVNVKDLTPAETDVAVAVKQWEEDMKESKETGIVNPSG 900
QY 901 TTVDKLVISKYKDFPFLASHHGVLGTSRPGCHTYVMYDDKGMQDEYVYKTYGLAFLSARC 960
Db 901 TTVDKLVISKYKDFPFLASHHGVLGTSRPGCHTYVMYDDKGMQDEYVYKTYGLAFLSARC 960
QY 961 RKPISIPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTREMEHFLOTNVKYPQMSPA 1020
Db 961 RKPISIPVPVHYAHLSCAKELRYTYKEHYIGDYAQPRTREMEHFLOTNVKYPQMSPA 1020
RESULT 2
Q60VS0_CAEBR PRELIMINARY; PRT; 976 AA.
AC Q60VS0;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
```

```
DE Hypothetical protein CBG19426.
GN Name=CBG19426;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI TaxID=6238;
RN [1]_
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF16;
RX PubMed=14624247; DOI=10.1371/journal.pbio.0000045;
RA Stein L.D., Bao Z., Blasiar D., Blumenthal T., Brent M.R., Chen N.,
RA Chinwalla A., Clarke L., Clee C., Coghlan A., Coulson A.,
RA D'Eustachio P., Fitch D.H.A., Fulton L.A., Fulton R.E.,
RA Griffiths-Jones S., Harris T.W., Hillier L.W., Kamath R.,
RA Kuwabara P.E., Mardis E.R., Marra M.A., Miner T.L., Minx P.,
RA Mullikin J.C., Plumb R.W., Rogers J., Schein J.E., Sohrmann M.,
RA Spieth J., Stajich J.E., Wei C., Willey D., Wilson R.K., Durbin R.,
RA Waterston R.H.;
RT "The genome sequence of Caenorhabditis briggsae: a platform for
RT comparative genomics."
RL PLOS Biol. 1:166-192(2003).
CC -!- CAUTION: the sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAC01000098; CAE72296.1; -; Genomic_DNA.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; P1wi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; P1wi; 1.
DR PROSITE; PS50822; P1wi; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 976 AA; 114611 MW; 8B5427451C9ADCAA CRC64;
SQ
Query Match 60.6%; Score 3241; DB 2; Length 976;
Best Local Similarity 62.7%; Pred No. 7 4e-174;
Matches 630; Conservative 136; Mismatches 206; Indels 32; Gaps 8;
QY 21 MKWLARPTGKCDGPKYKKVLLVNMVFPSSKIYDREYVEYVMTKE--VLNRKPKGKPF 78
Db 1 MKWLSRPEKCEGKYVVEKVKLLVNMVFNFTNIYDREYV-YFDMYKCKYMDKRTKSPK 59
QY 79 PKKTEIPDRAKLPWOHLRHEKKQTDIFLEDYVDFSKDTVYVCL-NTVTSKMLVSEK 137
Db 60 TKQMPYIPERSEAFWKHLRFEQKNPFQIEDYVDFSKDTVYSIHKTNGMSSEI- 114
QY 138 VVKQSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDEEANRSYKFLKNVMTQKV 197
Db 115 ---PDQNRDILKVLK-----MHYCKENLYFSREDPTRDEANRSCKFLKNVMTQKV 164
QY 198 RYAPFVNEIKVQPAKNFYDNNSTILRVPSFHDNRPEQSLEVPRIEAMFGIYIGKE 257
Db 165 RCAPVANEIGQLAKNFIYDGNSTILRVPSFHDNRPEHSLIEIAPRIESWFGIYIAVEK 224
QY 258 LFDEGPEVLNFAVDKLFYNAPKMSLDYLLIVDPSQNDVVRKDLTKLMAGKMTIROA 317
Db 225 LFDGEPVLNFAIDKLFYNAPQMSLDYLLIVDPSQNDVVRNRHKEQLNGSLTIKPA 284
QY 318 ARPRIQLLENLKLCAEVDNEMSRLETHERLTFDLDCEENSLYKVTGKSDRGNAKY 377
Db 295 MFQSDQLLANLKLKAASVWNPKRSEFAERHLLTFIRLSKYNSHEEMIIPVRGRDRNAPID 344
QY 378 DTTLFKIYEENKFTIEFPHLPLVKYKSAKEYAVPMHEHLEVEHVKPORYKNRIDLVMDQK 437
Db 345 RVPLFOIYEKRNKYIEFPRLPLVCKSGNNEYSVPMEFLEVYVEKPYKYNRIDFAMQDKF 404
QY 438 LKRATRKPDPHYKENTLKMKELDPSSEELNFVEREGLCSKQMIECPGKVLKEPMLVNSV 497
Db 405 VNAATRDPMHYKKEALEMLKDLDFSSGSLNFVERFGFTDLDKMI ECVGKVLKEPMLVNSV 464
```

```
QY 498 NEOIKMTPTVIRGFOEKQLNVVPEKELCCAVFVNE-TAGNPCLEENDVVKFYTELIGGCK 556
Db 465 NOKISMTFVVRGFOEKALNVVPEKELCCAVFLVRDEKTEPCLEEDVRFRFYKTLIDGCE 524
QY 557 FRGIRIGANENRGAQIMVDATKNBYAFYKNCYTLNTGIGRFEIATAEAKMFERLPDKEQ 616
Db 525 FRSIRIGKHONSQVRSLLQDPESGKYGFYVNVQLTAGVRNFRACANDAKAMFELADKNN 584
QY 617 KULMFIISKRQLNAVGVFKHYCDHTIGVANQHITSETVTYKALASLRHEKSGKRIFYQIA 676
Db 585 KILLFIVFSKRRWNFYGVKQFCDVBLGVASQHTVAFVAKRALQDMSSHKPKSKRIFYQIA 644
QY 677 LKINAKLGGINQELSDSETABISPEEKERKMTPLMTYTGIDVTHPTSGIDYSTAAVV 736
Db 645 LKINGKLGGVQNLQDSENAEMTVEBEKKKNMPLRMVYGVIDVTHPTSGIDYSTAAIV 704
QY 737 ASINPGGTYRNMIVTQEECRPERAVAHGRETDILEAKFVKLLREFAEENNDRAPAH 796
Db 705 ASNPGGTYRNMIVTQEEENRPERPVAGRETDILEGKFKVLLRIFAENNENRIPSH 764
QY 797 VYRVDGVSSEMLRVSHDELRSLSKSVKQPMSEDEGPEPKYTFIVIOQRHNRLLRRM 856
Db 765 VYRVDGVSSEMLRVSHDELRSLSKAEVRFLEKRGKNEQIPYTFIVLQKRHKTRLPK 824
QY 857 EKDKPVNVKDLTPAETDVAVAVKWHEEDMKESKETGIVNPPSGITTVDKLIVSKYKDP 916
Db 825 ENERP-----EAAKRWDEELKASQNTGFGVNPSTGTTVDRTIVSKYKDPF 872
QY 917 LASHHGVLGTSRPGHYTVMYDDKGMDSODEVYKMTYGLAPLSARCRPISLPVPHYAHL 976
Db 873 LASHHGALGTSRPGYTVMHDDKMKYKDELYKMTYELSFLSARCRPISLPVPHYAHL 932
QY 977 CEKAKELYRXYKEHYIGDYAQPRTHEMEHFLQTNVVKYPMGSPA 1020
Db 933 CEKAKELYRCKNDGPAHPROARREDIEQYLQTNIOYPMGSPA 976
```

RESULT 3

```
AGOL ARATH
ID AGOL ARATH STANDARD; PRT; 997 AA.
AC Q9SJL3;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-2006, entry version 1.
DT 07-MAR-2006, entry version 29.
DE Argonaute-like protein At2g27880.
GN OrderedLocusNames=At2g27880; ORFNames=tl2.20;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Rongstad C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Unayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768 (1999).
CC -!- SIMILARITY: Belongs to the argonaute family.
CC -!- SIMILARITY: Contains 1 PAZ domain.
CC -!- SIMILARITY: Contains 1 Piwi domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
```

```
DR EMBL; AC006929; AAD21514.1; -; Genomic_DNA.
DR PIR; A84678; A84678.
DR GenomeReviews; CT485783_GR; AT2G27880.
DR TAIR; At2g27880; -.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
DR CHAIN 1 997 Argonaute-like protein At2g27880.
FT /FTId=PRO_0000194070.
FT DOMAIN 359 471 PAZ.
FT DOMAIN 638 958 Piwi.
FT SQ SEQUENCE 997 AA; 111135 MW; DA93A1446C422P31 CRC64;
Query Match 11.3%; Score 604.5; DB 1; Length 997;
Best Local Similarity 23.8%; Pred. No. 3 1e-25;
Matches 242; Conservative 164; Mismatches 353; Indels 259; Gaps 41;
QY 38 KKVLLLVNMFKSSKIYDREYIEYEVKMTKEVLNRKPKGPFPPKTEIPIPRAKLFWQHL 97
Db 158 KKVWVRAN--HFLVQVADRDLVHYDVSINPEVSK-----YKDSHLGGKSPAYD-GRK 220
QY 98 RHEKQKTDFILEDYFDEKDTVYVCRNLNTVTSKMLVSEKVKVKDSE--KKDEKDLKK 154
Db 191 -----TVARNVNMKLLVKN---YKDSHLGGKSPAYD-GRK 220
QY 155 ILVTM--ILTYRKPHLNFSENPEKDEEANRSYKE-LKNV-----MTOKVRYA 200
Db 221 SLTTAGLPFDSKEFVNVLAEGKADSGSKDRPFKVAVRNVTSTDLVQLQOFLDRKOREA 280
QY 201 PFVNEBIKQVAKNFVYDNNLSILRVPESEFHPNRFQSL-----EVAPRIEAWFGI 251
Db 281 PV--DTIQVLDVVLDRDKPSNDYVSVGRSP-----FHTSLGKDARDGRGELGDIEYWRGY 333
QY 252 YGIGELFDGEPVINFVVDKLFYNAPKMSLLDYLLIIVDPQSCNDVDRKLTKLMAKG 311
Db 334 FOSLR-LTQWGLSLNIDVSARSFY-----EPIVVTDFISKFLNIRDL--N 375
QY 312 MTIROAARPRIOLLENLKLKCAEVWDNEMSLRHLTLFDLCEENSLVYKVTGKSD-- 369
Db 376 RPLRDSRLKVKVLRLLKVLKLLH-WN-----CTKSA---KISGISLP 415
QY 370 ----RGRNAKDYDTLLFKIYEENKFP-IEFPHLPLVKVSKGAKAYAVPMHEHLEVHEK 424
Db 416 IRELRTLEDKSEKTVVQVFAEKYVNRVYQALPAIQTGSDDTRPVLYPMELCQIDE-QGR 474
QY 425 YKNRIDLVMQDKFLKATRKPHDYKENTLKMELDFSEELNPFVERFGLCSKLOMIECP 484
Db 475 YTKRLNEKQVTTALLKATCQRPDP-RENSIKNLVYKNNYNDLIS--KEFGMSVTTQLASIE 531
QY 485 GKVLKEPMLVNSVNEQIKMTPVIRG---FOEKALNVVPEKELCCAVFVNVNATAGNCPLE 540
Db 532 ARVLPPPMKLYHDSGKERKVNPRLGOWNIDKQWNGAKVTSWTCVSFSTRIDRGLP--- 588
QY 541 ENDVVKFYTELIGGCKPGRIGRIGANENRGAQSIMYDATKNEYAFYKNCYTLNTGIGRFEI 600
Db 589 ----QEFCKQLIGMCSVCKGM---EFKQPAIPFISCPPEH-----IEEA 625
QY 601 ATEAKMFERLPDKEOKVLMPIIISKRQLNAVGVFKHYCDHTIGVANQHITSETVTYKALA 660
Db 626 LLD---IHKRAPGLQ---LLVILPDVTGSGYKIKRICETELGIVSQCCQVQVKNL-- 676
QY 661 SLRHEKSGKRIFYQIAKLNKAKLGGINQELSDSETABISPEEKERKMTPL-----TMVY 715
Db 677 -----NKQYMNVAKLNKVTGGRNTVLN-----DAIRNRIPLIDRPTIIM 718
QY 716 GIDVTHPTSGIDYSTAAVVASIN-PGGTYRNMIVTQ---EE-----CRPGERAV 763
Db 719 GADVTHPTQGEDSSPSIAAVVASMDWPEINKYKGLVSAQAHRFEIIQDLVKLVQDPQRL 778
QY 764 AHGRETDILEAKFVKLLREFAEENNDRAPAHIVVTRDGVSDSEMLRVSHDELRSLSKSEV 823
```


Qy	153	KKILYTMILT	YRKKKPHL	NFSREN	PEKDEAN	RSYKFL	KNVMTQ	KRYAP	PFVNEE	IKVQA	212									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Db	128	VSIKWLAT	VSWMGL	HALVSQ	IPVPLES	VQALD	VAMRHLAS	--MRYTPV	----	175										
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Qy	213	KNFVYDNN	SIILRVPS	FHDNRF	EQSLEVA	PIEAWFG	IYIGIKEL	FDEGPV	LNAI	VDK	272									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Db	176	-----	GRSFPSP	PEGYHP	-----	LGGRVWF	FGHQSV	RAM	--WKMM	LNDVSAT	220									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Qy	273	LFYNAPKMS	LDYLLI	VDPOCND	VRKDLK	TKLMAG	MTIROAAR	PIRQ	LLLENL	KLK	332									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Db	221	AFYKA--	QPVI	EFCMEV	LDIRN	QPK-----	PLTDSQ	RVFR	TKIKGL	KVE	266									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::											
Qy	333	---CAE	VDN----	EMSRL	TERHLT	FLDL	CENSL	VYKV	TKSG	DRNNAK	YDTTL	FKIY	385							
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::									
Db	267	VTHCGM	KRYKRV	CNVTR	PPASHQ	TFPLQ	LESGQTV	-----	ECTVAQ	YFV	310									
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::										
Qy	386	EENKFP	--IEFPH	LPLV	KVKS	GAKEVA	PMWEH	LEVHE	QPVY	KNRIDL	VMQDK	FLKRA	YTK	444						
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::								
Db	311	KQKYNLQ	LKYPHL	PCLVQ	GOQK	HTYPLE	VNCNI--	VACOR	CIKKL	TNQTS	MTWIKAT	ARS	369							
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::									
Qy	445	PHDYKENT	LKML	KELDF	SSBELN	FVER	GLCS	KLQMI	BCPG	KVLK	PEMLV	NSVNB	QI	KMT	504					
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::								
Db	370	APDRQEB	ISRL	MKNAS	YNLDP--	YIQE	FGIK	VKDD	MTV	TRVL	PAPILQ	YVGR	NR	AIAT	427					
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::								
Qy	505	P-----	VIRG	QEKOL	--NV	PEKEL	CCAV	VVN	ETAG	NPCL	EEND	VVFY	TELI	-----	552					
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::							
Db	428	PNQGV	MDRG--	--KOFY	NGIB	EIKW	AIAC	PAPOK	-----	QCR	EE--	--VL	NFTD	QL	RSKD	478				
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::							
Qy	553	GGCKFR	GIRIG	ANENR	GAQS	TM	YDAT	KNEY	AFYK	NYKNT	LNTG	IGR	PIEA	ATEAK	NF	REL	612			
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::					
Db	479	AGMPI	QOQ	CFCK	YQAQ	ADSV-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	507			
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::					
Qy	613	DKEQK	VL	FIIS	KRL	NAYG	VKHY	CDHT	IGV	ANQ	HIT	SET	VT	VTKAL	ASLR	HEK	SKRIF	672		
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::				
Db	508	NTYSGL	QLII	VIL	PKTP	PVYAE	VKRV	GD	TLLG	MA	TQC	VQV	KNV	VKT-----	SPQT	L	558			
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::					
Qy	673	YQI	TALK	NAKL	GGIN	QEL	NDSE	IAEIS	PEEK	ERRK	TMPL	TMV	GV	GD	VTH	PT	SYSG	DI	YSI	732
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::	::: :::			
Db	559																			

RESULT 5
Q4RKH3 TETNG
ID Q4RKH3_TETNG PRELIMINARY; PRT; 863 AA.
AC Q4RKH3_
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 21 SCAR15029, whole genome shotgun sequence. (Fragment).
GN ORFNAMES=GSTENG00032955001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
CC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
CC	Tetraodontidae; Tetraodontidae; Tetraodon.
CC	NCBI_TaxID=99883;
RN	[1]
RN	NUCLEOTIDE SEQUENCE.
RX	PubMed=15496914; DOI=10.1038/nature03025;
RA	Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA	Maucel E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA	Nicaud S., Jaife D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA	Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA	Bisemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA	Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA	Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA	Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA	Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA	Winkler P., Lander E.S., Weissbach J., Roest Crallius H.;
RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT	the early vertebrate proto-karyotype.";
RL	Nature 431:946-957(2004).
RL	[2]
RP	NUCLEOTIDE SEQUENCE.
RG	Genoscope; Whitehead Institute Centre for Genome Research;
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC	-!- CAUTION: The sequence shown here is derived from an
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
CC	-----
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
DR	ENSL; CAARE0105029; CAG11109.1; -; Genomic_DNA.
DR	SNR; Q4RKH3; 206-373.
DR	InterPro; IPR003100; PAZ.
DR	InterPro; IPR003165; Piwi.
DR	Pfam; PF021170; PAZ; 1.
DR	Pfam; PF021171; Piwi; 1.
DR	PROSITE; PS50821; PAZ; 1.
DR	PROSITE; PS50822; PIWI; 1.
FT	NON_TER 1 1
FT	NON_TER 863 863
SQ	SEQUENCE 863 AA; 98064 MW; 8D4C8A0741D8E183 CRC64;

Query Match	10.7%	Score	573	DB 2	Length	863			
Best Local Similarity	23.2%	Pred.	No. 1.5e-23						
Matches	240	Conservative	176	Mismatches	349	Indels	268	Gaps	46

QY	19	PEMKWLARP	GTG	CDG	KFY	EKKV	YKVV	LLVNW	PF	SSKI	YDB	REY	YEEY	YVQMT	KEV	LN	RPG	KPP	78																																							
Db	1	PVPEYV	KPPSR	PD	FGT	MG	RTIK	LQ	AN	FFEM	--EIP	KL	EV	YH	YD	DI	-----	KPEK	-C 50																																							
QY	79	PKKTEI	PI	PR	AKL	FM	QHL	RHE	KQ	TDF	ILE	DY	VD	ED	KD	TV	VS	CR	LT	NT	VT	SK	ML	SV	SE	KV	138																															
Db	51	PRVNR	IE	VE	HM	-----	VQ	H	P	KTQ	-----	IF	G	R	K	P	YD	-----									79																															
QY	139	VKKD	SEK	KD	KE	LE	K	IL	Y	TW	--	IL	T	Y	R	K	P	H	L	N	F	S	R	EN	P	K	D	BE	AN	S	Y	K	L	N	V	Q	196																					
Db	80	-----	GR	KL	Y	T	AM	P	I	G	R	K	D	K	V	E	L	T	I	P	G	E	K	D	R	S	F	K	V	S	I	K	V	S	C	L	Q	126																				
QY	197	VR	Y	A	-----	P	V	N	E	E	I	K	V	Q	A	K	N	F	Y	D	N	S	T	L	R	---	V	P	E	S	H	D	N	R	F	Q	S	L	E	V	A	P	R	245														
Db	127	LH	E	A	L	S	R	L	S	V	P	F	--	E	T	Q	---	A	L	D	V	M	R	H	L	P	S	M	R	Y	T	P	G	R	S	F	T	P	G	G	C	A	N	P	L	G	G	G	R	180								
QY	246	E	A	W	F	G	I	Y	I	G	I	K	E	L	P	D	G	S	P	V	L	N	F	A	I	D	X	L	F	N	A	P	K	M	S	L	L	D	Y	L	L	I	V	D	P	O	S	C	N	D	D	V	R	K	O	L	T	305
Db	181	E	V	N	F	G	H	Q	S	V	R	P	S	L	-	W	K	M	L	N	I	D	V	S	A	T	A	F	Y	K	A	---	Q	P	V	I	E	F	C	E	V	L	D	F	K	S	I	E	E	Q	-----	230						
QY	306	K	L	M	A	G	K	T	I	Q	A	R	P	R	I	Q	L	L	E	N	K	L	K	---	C	A	E	V	D	N	---	E	M	S	R	L	T	E	R	H	T	I	F	L	D	C	S	E	N	358								
Db	231	-----	O	K	P	L	D	S	O	R	V	K	E	T	K	E	I	G	L	K	V	I	T	H	C	G	O	M	K	R	Y	V	C	N	V	T	R	R	A	S	H	O	T	F	-	P	L	O	E	N	284							

Qy	197	QY	RYA	-----	-P	VNEEIKVQFAKNFYDNNISILR	---	VPESHDPNRFQSLSEVAPRI	24
Db	127	LHEAL	SGLRSPVFP	--ETVQ	---	ALDVVMRHLPSMRVTPVGRSGFTPTSPGCGANPLGGGR	---		180
Qy	246	EAMFG	IYIGIKELPDGPRVLNFAIVDKL	FYNAPKMSLLDYLILLIV	DPSCNDVDVKDLKT	305			
Db	181	EWVGF	HQSVRPSL	-WKMM	INDVSATAPYA	---	QPVIEFMCVELDFKISIEQ	-----	230
Qy	306	KLMAG	KWTTIROAARPR	IQLLENLKLK	---	CAEVMDN	---	EMSLRTEHHTFLDLCBEN	358
Db	231	----	OKPLDTSORVKETKIGLKVITHCGOMKRYRVGVNVTTRPASHOTF	---	PLOOEN	284			

DT 21-FEB-2006, entry version 11.
 DE Putative leaf development and shoot apical meristem regulating
 DE protein.
 GN Name=OSJNBb0070009.6;
 OS Oryza sativa (japonica cultivar-group).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsirlin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.I., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBb0070009 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; AC087412; AAP68386.1; -; Genomic DNA.
 DR Gramene; Q7Y001; -;
 DR InterPro; IPR003100; PAZ.
 DR InterPro; IPR003165; P1wi.
 DR Pfam; PF02170; PAZ; 1.
 DR Pfam; PF02171; P1wi; 1.
 DR PROSITE; PS0821; PAZ; 1.
 DR PROSITE; PS0822; P1wi; 1.
 SQ SEQUENCE 1055 AA; 115887 MW; 1817D71A0B71B3DD CRC64;

Query Match 10.6%; Score 568.5; DB 2; Length 1055;

Best Local Similarity 22.4%; Pred. No. 3.6e-23;
 Matches 226; Conservative 169; Mismatches 345; Indels 271; Gaps 39;

QY 49 FSKYIDREYVEYEVKMTKEVLRNPKGPFPPKTEIPIDRAKLEFQHLRHEKKTQDFIL 108
 DB 213 FLVQADKDIHYDVVITPESTYRNRISIINKL-----VALHKQFLDGLR 258
 QY 109 EYVDEKDTVTSVCRNLNTVTSKMLSEKVKVDEKDE-----KLEKKILYIMILTY 163
 DB 259 P--VYDGRKSIYTAGPLPKTKDFVVGKHNPLRGNQRESEYKVTIKQASKTDLYSL---- 312
 QY 164 RKKEHLNFRSRENPEKDEBANRSYKFLKNVMTQKRVYAPVFNBEIKVQPAKNFVYDNNSTL 223
 DB 313 -KQFLVGRQRELPQDTIQA-----LDIALRECP-----TSVNFCTDR--YV 350
 QY 224 RVPESFHPDNRPEQSL-----EVAPRIEAWFGIY-----IGIKEL 258
 DB 351 SISRSE-----FSQFSGHGEIGSGTECWRYGYSRLPTQGLSLNIGMDLPQNISATAF 405
 QY 259 FQGEVPLNPAIVDKLPYNAPKMSLLDYLIIVDPOSCDDVDKDKTKLMAGMTIROQA 318
 DB 406 YKAQPMVMPAV-----QYLN-----RDVSRRL-----SDQD 432
 QY 319 RPRIRQLLENLAKCAEVWDNEMSRITRHLTFLDLCEENSLVYKVTGKSDRGRNAKYD 378
 DB 433 RIKLKALKGVQI-VATHW-----KEKSYRIKITGIPSPAMNELMPD 473
 QY 379 TT-----LFKYIYENKKEFIEFPHLPVVKVSGAKAYAVPMHELVHEKPYQXRI 429
 DB 474 LOGNRISVVQYFKQYNSLKHVNPCL---QAGSDSRPKYLPMEVCSILE-QGRYSKK 529
 QY 430 DLYMQDKFLKRAIRKPHDYKENTLKMELDFSSSELNFRFGLSKLQMTCEGKVLK 489
 DB 530 NBEHQVTNLRMTCPERPAQRESSIIIEVNTNSYGNDDC--AKEFGIKVANQLAVVDARVLP 587

QY 490 EMLV--NSVNBQIKMTVPVIRGQEQKQLNVVPEKELCCAVFVNVNAGN--PCL----- 539
 DB 588 TPLRYKHSGRKV--CNFSV-----GQNMNINR-----MVNGGCIHNHTCLSFASRM 634
 QY 540 EENDVVKYETELIGGCKFRGIRIGANENRGAOSIMYDATKKNFYAFYKNTCLNTLTGIGRFEI 599
 DB 635 HVNDIRMFCEDLVGMNNGMGM---NTRPCVDIIIGQQORNIETGAIRNI----- 680
 QY 600 AATEAKNMFERLPDKE---QKVLMIISIKROLNAYGVKHYCDHTIGVANQHITSETVT 656
 DB 681 -----HROSSEKLDQDQDLTGQQLQLLVILTETSGSYGRIKRICETEVGV-----IT 727
 QY 657 KALASLRHEKSGKRIFYOALKINAKINAGGINOELOWSEIAETSPKEKRRKTMPL----- 711
 DB 728 OCCAPKSKQKQKQVLENLAKNNVKGGRNTVL-----EDALHKKIPILTRDP 776
 QY 712 TMYGIDVTHPTSYSGIDYSIAAVASIN--PGGTIYRNMIVTQE-----ECRPG 759
 DB 777 TIVFGADVTHPSPGEDASPSIAAVASMDWPEVTYKCLVSTQSHREEIISNLYTEVKDP 836
 QY 760 ERAVAHGRERTDILEAKFVKLLREFARNNDNPAHIIVYRDGVSDEMLRVSHDELRLSL 819
 DB 837 LKGIIRG-----GMIRELLRSFYQETGQK-PSRIIFPYRDGISEGQFSQVLLYEMDAI 887
 QY 820 KSEVKQPMSEKDEGDEPEPKYTFIVIQKRNTRLRLRMEKDKPVVKNKDLTPAETDVAVAV 879
 DB 888 R---KACASLQBGY--LPVTFVVKRHHHTLFPENRDMDRSGNIIIP----- 932
 QY 880 QWMEEDMKESKETGIVNPSGTTVDKLVSKYKPDFFLASHHGVLTGSRPHYTVMYDDK 939
 DB 933 -----GTVDVTMICHPSSEDFYLCSHSGIKGTSRPHYHVLLEN 972
 QY 940 GMSQDEVVMTYGLAFLSARCKPISLPVPHYAHLSCEKAKELYRTYKEH 990
 DB 973 GFKADTLQTLTYNLSYTYARCTRAVSIVPPAYYAHLAGAPRAR--YMEDEH 1021

RESULT 9
 QPZ41 ANOGA
 ID Q7PZ41 ANOGA PRELIMINARY; PRT; 891 AA.
 AC Q7PZ41;
 DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
 DT 07-DEC-2004, sequence version 2.
 DT 07-FEB-2006, entry version 8.
 DE ENSANGP0000008896 (Fragment).
 GN ORFName=ENSANGG0000008700;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 OC Anophelinae; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; AAA801008986; EAA00062.2; -; Genomic DNA.
 DR SMR; Q7PZ41; 239-406.
 DR InterPro; IPR003100; PAZ.
 DR InterPro; IPR003165; P1wi.


```
QY 49 FSSKIYDREYVEYVQMTKEVNLNRKPGKPPKTEIPIDRAKLFQWHLRHEKKQDTFTL 108
Db 47 FEMDTPKIBIYHDDI-----KPEK-CPRVRNREIVHEM-----VOHFAQ---IF 89
QY 109 EYD--VFDEKOTVYVCRNLNTVTSKMLVSEKVKDKSEKKDLEKKILYTMILTYRKK 166
Db 90 GDRKPVDFGRKNLYTAMPLPIARDKQV--ELEVTLPGCKDR--IFKVAIKMACVSLQA 145
QY 167 FHLNFSRENPEKDEENRSYKFLKNVMTQ--KVRVAPFVNEIKVQFAKNFYDNNLSILR 224
Db 146 LHDALSGRLPNVPFE---TVQALDVVMRHLPSMRYTPV-----GRSFFT 186
QY 225 VPSEFHPDNRPFQSLSEVAPRIEAWFGIYIGIKELFDPGEVPLNFAIVDKLFYNAPKMSLLD 284
Db 187 ASEGCANP-----LGGREVGWFGFHSVRPSL-WKWMNLNIDSATFYKA--QPVE 235
QY 285 YLLLVDPQSCNDVVRKDLTKMLAGAKMTIROAARPRIRQLLENKJLK---CAEVDN-- 339
Db 236 FMCEVLDFKSIIEQ-----QKPLTDSQRVKFTKEIKGLKVEITHCGQMKRYR 283
QY 340 --EMSLRTERHLTFLDCEENSLVYKVTCKSDRGNAKKYDTTLPKIYEENKKFIEFPHL 397
Db 284 VCNVTRRPASHQTF-PLQESGQTVECT-----VAQY-----FK--DRHKLVLRYPHL 328
QY 398 PLVKVKSAGKAYAVPMLEHVEHEKQRYVKNRIDVMQDKFLKRAT-RKPHDYKENTLKM 456
Db 329 PCLQVGOEQKHLYPLEVCNIVAGORCIKLLTD--NOTSTMIRATARSAPDQBEISKUM 386
QY 457 KELDFSSEELNVEFRGLCSKLOMECPGKVLKEPMLVNSNEQIKMTVP--IRGFQBEQ 514
Db 387 RSASENTDP--FVREFGIMVKDDMTDVTGRVLQPPSILYGRSKAIAIPVQGVMDNRNQ 444
QY 515 LNV-VPEKELCAVFNWETAGNPLEENDVVKFTTELI-----GGCKFERGIRIGANEN 567
Db 445 FHTGTEIKVWAIACFAPO---RQCTEVH--LKTFTQLRKISRDAGMPIQGPCCKYA 498
QY 568 RGAQSI--MYDATKNEYAFYKNCNTLNTGIGRFEIAATAKNMFERLPDKEQKVLFAIIS 625
Db 499 QADSVPEMFRHLKTY-----TGL-----QLVVIL 525
QY 626 KQLNAYGVFVGHYCDHTIGVANQHTSETVTTKALASLRHEKSGKRIFYOIALKINAKLGG 685
Db 526 PGKTPVYAEVKRVGDTVLMGATQCQVMKNV-----QRTTPQTLNLCIKINVKLGG 576
QY 686 INQELDWSIAISIEPEEKERRKTMPLTMVYGVGDVTHPTSYSIGIDYSIAAVASINPGGTI 745
Db 577 VNNIL-----LPQGRPPVFQPV-IFLGADVTHPPAGDGKPKSIAAVGMSMDAHPNR 627
QY 746 YRNMTVTOBECRGERAVAHGRERTDILEAKFVKLLRBFPAENNDNRAPAHIVVYRDGVSD 805
Db 628 YCATVRVQ-----HROEIIQDLSAMVRELLIQYKSTRPK-PTRIIIFYRDGVSE 676
QY 806 SMLAVSHDELASLSEKVKQFMSERDGEDPEPKYTFIVIQKRHTLLRRMEKDKPVNKK 865
Db 677 GQFQVQLVHELLAIREACIKL-----EKDYQPGITFIVVQKRHTLRF-CTDRNERVGKS 730
QY 866 DLTRAETDVAVAVKQWEDMKESKETGIVNFSSTGTVDKLIVSKYKDFPLASHGVLG 925
Db 731 GNIPA-----GTTVDTKITHSEDFDYLCSHAGIQG 761
QY 926 TSPRGHYTMVDDKMGMSQDEVYKMTYGLAFLSARCKRKPISLPVPVHYAHLSEKAKELYR 985
Db 762 TSPRSHYVLDNDRFSSDELQILYQLCHTVYVTRTSVSIIPAYAYAHVAFRAR-YHL 820
QY 986 TYKEH 990
Db 821 VDKEH 825
RESULT 11
12C2 RABIT
ID 12C2 RABIT
AC 077503; PRT; 840 AA.
```

```
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 14-NOV-2003, sequence version 2.
DT 07-FEB-2006, entry version 30.
DE Eukaryotic translation initiation factor 2C 2 (eIF2C 2) (eIF-2C 2)
DE (Argonaute-2) (Fragment).
GN Name=eIF2C2; Synonyms=AGO2;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OC Oryctolagus.
NCBI_TaxID=9986;
[1]
RN NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.
RP TISSUE=Liver;
RC MEDLINE=98267198; PubMed=9602122; DOI=10.1016/S0378-1119(98)00107-3;
RA Zou C., Zhang Z., Wu S., Osterman J.C.;
RT "Molecular cloning and characterization of a rabbit eIF2C protein.";
RL Gene 211:187-194(1998).
CC -!- FUNCTION: Provides endonuclease activity to RNA-induced silencing
CC complexes (RISC). Cleaves siRNA/mRNA heteroduplexes bound to RISC.
CC Essential for embryonic development as well as RNA-mediated gene
CC silencing (RNAi) (By similarity).
CC -!- SUBUNIT: Interacts with DICER1 through its Piwi domain (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic bodies (By similarity).
CC -!- DOMAIN: Piwi domain is thought to provide RNA cleavage activity
CC with a mechanism similar to RNase H (By similarity).
CC -!- SIMILARITY: Belongs to the argonaute family.
CC -!- SIMILARITY: Contains 1 PAZ domain.
CC -!- SIMILARITY: Contains 1 Piwi domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF005355; AAC24323.1; ALT_INIT; mRNA.
CC
DR PIR; PC6505; JC6569.
DR SMR; O77503; 203-370.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
KW Direct protein sequencing; Initiation factor; Protein biosynthesis.
FT CHAIN <1 840
FT Eukaryotic translation initiation factor
FT 2C 2.
FT /FTID=PRO_0000194059.
FT PAZ.
FT DOMAIN 216 329
FT DOMAIN 498 799
FT NON_TER 1
FT SEQUENCE 840 AA; 95306 MW; 1E703F9E31391F29 CRC64;
SQ
Query Match 10.6%; Score 565; DB 1; Length 840;
Best Local Similarity 23.2%; Pred. No. 4.1e-23;
Matches 231; Conservative 157; Mismatches 357; Indels 250; Gaps 39;
QY 38 KVVLLLVNMFESSKIYDREYVEYVQMTKEVNLNRKPGKPPKTEIPIDRAKLFQWHL 97
Db 17 RTIKLQANFFEMDIPKIDYHIELDKPEK-----CPRVRNREIVHEM-----V 60
QY 98 RHEKKQDTFTILEDY--VFDEKOTVYVCRNLNTVTSKMLVSEKVKDKSEKKDLEKKI 155
Db 61 QHFAQ---IFGDRKPVDFGRKNLYTAMPLPIGREKV--ELEVTLPGCKDR--IFKVS 112
QY 156 LYTMILTYRKKFHLNFSRENPEKDEENRSYKFLKNVMTQ--KVRVAPFVNEIKVQFAK 213
Db 113 IKWVSCVSLQALHDALSGLPSVPFE---TIQALDVVMRHLPSMRYTPV----- 158
QY 214 NFVYDNNISILRVPSFHPDNRFEQSLSEVAPRIEAWFGIYIGIKELFDPGEVPLNFAIVDKL 273
Db 159 -----GRSFFTASEGCSNP-----LGGREVGWFGFHSVRPSL-WKWMNLNIDSATA 204
QY 274 FYNAPKMSLLDYLILLIVDPQSCNDVVRKDLTKMLAGAKMTIROAARPRIRQLLENKJLK- 332
```

Db 205 FYKA ---QPVIEFVCEVDFKSIHQ-----QKPLTDSQRVKFTKRIKGLKVEI 250
Qy 333 --CAEVDN---EMSRLTHERLFTDLCEBNSLVYKVTGKSGRGRNAKYDTTLFKIYE 386
Db 251 THCQMKRKRYKVCNVRPASHQTF-FLQGESGQTVECT-----VAQY---FK--D 295
Qy 387 ENKFIPIPHPLPLVKVKGAKYAVPMHLEVEHKPORYKVRDLVMDQKFLKRAT-RKP 445
Db 296 RHKLVLRYPHPLCLQVGOEQRHTYPLLEVNCNVAGQRCIKGLTD--NOTSTMTRATARS 353
Qy 446 HDYKENTLKLMLKELDFSEBELFVERFGLCSKLQMIKPCGKVLKPEKPMVNSVNEIQIWT 505
Db 354 PDQEBTSLKRSASFNDP--YVREFGIVKIDEMTDVTCRVLPQPPSILYGGRNKATAP 411
Qy 506 V--IRGPOEQQLNV-----VPEKELCAVFEVNET-----AGNP 537
Db 412 VQGVMDNRNQFHTGIBIKVWATACFAPQKQ-CTEVLHKSFTQELRKISRDAAGMPIOQ 470
Qy 538 CLEENDVVKYTELIGGCKEGRIGRIGANENRGAQSI--MYDATKNEYAFYKNCITLNTGIG 595
Db 471 CF-----CKYA-----QGADSVGPMFRHLKNTYA-----494
Qy 596 RPEIAATEAKNMFELPDKEQKVLMIISKQLNAYGFVKHYCDHTIGVANQHIITSETV 655
Db 495 -----GLQVVVILPGKTPVYAEVRVGDVTLGMAVCQVMKNV 533
Qy 656 TKALASLRHKGSKRIPYQIALKINAKLNGINQELDMSIABISPEKERRKTMPLTMYV 715
Db 534 -----QRTTPQTLNCLKINVLKGGVNL-----LPQGRPPVFPQPV-IPL 575
Qy 716 GIDVTHPTSGDIDYSTAAVAINPGTIRNMIVTQBCRGERAVAGHRTDILEA 775
Db 576 GADVTHPPAGDGKKPSTAAVVGSDAHNPNYCATRVQQ-----HROBIIQDLAA 625
Qy 776 KFKVLLREFAEENNDNRAPAHIVVYRQVSDSEMLRVSHDELRLSKRVKQFMSEKDEDP 835
Db 626 MVRELLIQFYKSTRFK-PTRIIFRQVDSGQVQOVLHLELLAIREACIKL-----EKDY 679
Qy 836 EPKYTFIVTQKRNTRLRLRMEKDKPVVKNKDLTPAETDVAAVAKQWEEDMKSKEGTIV 895
Db 680 QPGITFIVQKRRHRLF-CTDKNERVKGSGNIPA-----713
Qy 896 NPSSGTTVDKLVSKYKDFEFLASHGVLTSPGHYTVYDDKMSQSDRVYKMTVGLAF 955
Db 714 ----GTVDTKITHPTFEDFYLCSHAGIQGTSRPSHYVLMDDNRFSSDEBLQILTYQLCH 769
Qy 956 LSARCKRPISLPVPVHYAHLSCBKAKEYLRYTKEH 990
Db 770 TYVRCRVSIPAPAYAHVAFRAR-YHLVDKEH 803

RESULT 12
I2C2 MOUSE

AC Q8CJG0; STANDARD; PRT; 860 AA.
DT 01-NOV-2003, integrated into UniProtKB/Swiss-Prot.
DT 07-FEB-2006, entry version 18.
DE Eukaryotic translation initiation factor 2C 2 (eIF2C 2) (eIF-2C 2)
DE (Argonaute-2) (Slicer protein) (Piwi/argonaute family protein
DE meIF2C2).
GN Name=eIF2c2; Synonyms=Ago2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX PubMed=12526743; DOI=10.1016/S0960-9822(02)01394-5;
RA Doi N., Zengo S., Ueda R., Ohki-Hamazaki H., Ui-Tei K., Saigo K.;
RT "Short-interfering-RNA-mediated gene silencing in mammalian cells

RT requires Dicer and eIF2C translation initiation factors.";
RL Curr. Biol. 13:41-46(2003). [2]
RN FUNCTION, TISSUE SPECIFICITY, NULL MUTANT, AND MUTAGENESIS OF LEU-141;
RP ASP-598; GLN-634; HIS-635; ASP-670; HIS-683; PHE-705 AND THR-745.
RX PubMed=15284456; DOI=10.1126/science.1102513;
RA Liu J., Carmell M.A., Rivas F.V., Marsden C.G., Thomson J.M.,
RA Song J.-J., Hammond S.M., Joshua-Tor L., Hannon G.J.;
RT "Argonaute2 is the catalytic engine of mammalian RNAi.";
RL Science 305:1437-1441(2004).
CC -!- FUNCTION: Provides endonuclease activity to RNA-induced silencing
CC complexes (RISC). Cleaves siRNA/mRNA heteroduplexes bound to RISC.
CC Essential for embryonic development as well as RNA-mediated gene
CC silencing (RNAi).
CC -!- SUBUNIT: Interacts with DICER1 through its Piwi domain (By
CC similarity).
CC -!- INTERACTION:
CC Q9UPV3:DICER1 (xeno); NbExp-2; IntAct=EBI-528299, EBI-395506;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic bodies (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitous expression in 9.5 day embryos with
CC highest levels in forebrain, heart, limb buds, and branchial
CC arches.
CC -!- DOMAIN: Piwi domain is thought to provide RNA cleavage activity
CC with a mechanism similar to RNase H.
CC -!- MISCELLANEOUS: Mice lacking Eif2c2 are embryonic lethal with a
CC strong defect in neural tube closure and apparent cardiac failure.
CC -!- SIMILARITY: Belongs to the argonaute family.
CC -!- SIMILARITY: Contains 1 PAZ domain.
CC -!- SIMILARITY: Contains 1 Piwi domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB081472; BAC15767.1; -; mRNA.
DR SWR; Q8CJG0; 223-390.
DR IntAct; Q8CJG0; -
DR Ensembl; ENSMUSG00000036698; Mus musculus.
DR MGI; MGI:2446632; Eif2c2.
DR GO; GO:0005515; P:protein binding; IPI.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS00821; PAZ; 1.
DR PROSITE; PS00822; PIWI; 1.
DR Initiation factor; Protein biosynthesis.
KW Initiation factor; Eukaryotic translation initiation factor
FT CHAIN 1 860
FT 2C 2
FT /FTID=PRO_0000194058.
FT PAZ.
FT DOMAIN 236 349
FT DOMAIN 518 819
FT MUTAGEN 141 141 L->W: No effect.
FT MUTAGEN 598 598 D->A: Loss of RNA cleavage. Binds siRNA.
FT MUTAGEN 634 634 Q->A: No effect.
FT MUTAGEN 634 634 Q->R: Loss of RNA cleavage. Binds siRNA.
FT MUTAGEN 635 635 H->P/A: Loss of RNA cleavage. Binds
FT siRNA.
FT MUTAGEN 670 670 D->A: Loss of RNA cleavage. Binds siRNA.
FT MUTAGEN 683 683 H->Y: No effect.
FT MUTAGEN 705 705 F->Y: No effect.
FT MUTAGEN 745 745 T->Y: No effect.
SQ SEQUENCE 860 AA; 97294 MW; DA32A759EA467B31 CRC64;
Query Match 10.6%; Score 565; DB 1; Length 860;
Best Local Similarity 23.2%; Pred. No. 4.2e-23;
Matches 231; Conservative 159; Mismatches 355; Indels 250; Gaps 40;
Qy 38 KKVLLLVNWFKSSKIYDEYEVKMTKEVNLNRKPGKFPFKTIPIDRAKLFQWHL 97
Db 37 RTIKLQANFFEMDIPKIDYHYELDIKPEKR-----PRVNRIVEH-----V 80
Qy 98 RHEKQTDILEDY--VPDEKTVISVCLNTVTSKMLVSEKVKVKKDSEKKDKLEKKI 155


```

Db      254 TDSQVKTKEIKGLKVEITHCGQMKRYKRVNVTTRPASHQTF-PLQESGQTVECT-- 310
QY      368 SDRGRNKKYDVTTLFKIYENKFIPIPLVPLVKVSGAKAYAVPMHLEHVEHPQRYKN 427
Db      311 -----VAQY-----FK--DRHKLVLRYPHLPCLQVQEQEQRHTYLPLEVCNIVAGQRCIKK 358
QY      428 RIDLVMQDKFLKRAT-RKPHDYKENTLKLKELDFSSSEELNFVERFGKSLQWIECPGK 486
Db      359 LTD--NOTSTMIRATARSADPRQIEISKLMSASFTNDP--FVREFGIMVKKDDMTDVTGR 414
QY      487 VLKEPMLV-----NSVNEQIKMTFV--IRGFQEKQLNV-VPEKELCCAVFVNNETA 534
Db      415 VLQPPSILYGRVWEENAPLNKAIAITPVQGVMDNRNKQHTGTEIKVWALACFAPQ-- 471
QY      535 GNFCLRENDVVKFTTELI-----GGCKFRGIRIGANENRGAOSI--MYDATKNYAFYK 586
Db      472 -RQCTEVH--LKTFTQRLKISRISDAGMPIQCPFCFYAQCADSVPMFRLKNTY----- 524
QY      587 NCTLNTGIGRFEIATAEAKMFERLPDKEQVLMFIIISKQLNAYGVKHYCDHTIGVA 646
Db      525 -----TGL-----QVUVVILPGKTPVVAEVRKVGDTVLGMA 555
QY      647 NOHITSETVTKALASLREHSGKRIFYQIALKINAKLGINOBLDWSIEAIEISPEEKERR 706
Db      556 TQCVQMKNV-----QRTTPTLSNLCLKINVLGGVNNIL-----LPQGRPPV 598
QY      707 KTWPLTWYGVIDYTHPTSYSGIDYSIAAVASINPGGTIYRNMIVTQECRPGERAVAHG 766
Db      599 FOQPV--IFLGADYVTHPPAGDGKPSIAAUVGSDMAHPNRYCATVRVQQ-----HR 647
QY      767 RERTDILEAKFVKLLREFAEANNDRAPAHIVVVRDGVSDSEMLRVSHDELRLSKSEVKQF 826
Db      648 QEIQLDSAMVRELLIQFYKSTRFK-PTRIIFYRDGVSEGOFOVLHHELLAIREACIKL 706
QY      827 MSRBDEGDEPPKTYFVIQKRHNTRLIRREKDKPVVVKDLTPAETDVAVAVKQWQEDM 886
Db      707 -----EKDYQPGITFIVVQKREHTRLE-CTDRNVRKSGNIFA----- 744
QY      887 KESKFTGIVNPPSGTVDKLVISKYKPDFPLASHGVGLTSRPGHYVTMYDDKMGSDQEV 946
Db      745 -----GTTVDTKITHPSBDFLYLCSHAGIQGTSRPSHYVLMWDDNRFSDEL 791
QY      947 YKMTYGLAFLSARCRPISLPVHVHVLSCAKELRYRYKEH 990
Db      792 QILTYQLCHTYVRCRVSVPAPAYYAHVAFR-YYHLVDKSH 834

RESULT 14
ID_12C2 RAT STANDARD; PRT; 860 AA.
AC Q9QZB1;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 14-NOV-2003, sequence version 2.
DT 07-FEB-2006, entry version 26.
DE Eukaryotic translation initiation factor 2C 2 (eIF2C 2) (eIF-2C 2)
DE (Argonaute-2) (Golgi ER protein 95 kDa) (GERP95).
GN Name:EIF2c2; Synonyms:Ago2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Sprague-Dawley; TISSUE=Hepatoma;
RX MEDLINE=99443791; PubMed=10512872;
RA Cikaluk D.E., Tahbaz N., Hendricks L.C., DiMattia G.E., Hansen D.,
RA Pilgrim D., Hobman T.C.;
RT "GERP95, a membrane-associated protein that belongs to a family of
RT proteins involved in stem cell differentiation.";
RL Mol. Biol. Cell 10:3357-3372(1999).
CC -!- FUNCTION: Provides endonuclease activity to RNA-induced silencing
CC complexes (RISC). Cleaves siRNA/mRNA heteroduplexes bound to RISC.

```

```

CC      Essential for embryonic development as well as RNA-mediated gene
CC      silencing (RNAi) (By similarity).
CC      -!- SUBUNIT: Interacts with DICER1 through its Piwi domain (By
CC      similarity).
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic bodies (By similarity).
CC      -!- DOMAIN: Piwi domain is thought to provide RNA cleavage activity
CC      with a mechanism similar to RNase H (By similarity).
CC      -!- SIMILARITY: Belongs to the argonaute family.
CC      -!- SIMILARITY: Contains 1 PAZ domain.
CC      -!- CAUTION: Was originally (Ref.1) thought to be membrane-associated.
CC
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC      EMBL: AF195534; AAF12800.1; ALT_INIT; mRNA.
CC      Ensembl: ENSRNOG0000008533; Rattus norvegicus.
CC      SRR: Q9QZB1; 223-390.
CC      RGD: 621255; Elf2c2.
CC      GO: 0030154; P:cell differentiation; NAS.
CC      InterPro: IPR003100; PAZ.
CC      InterPro: IPR003165; Piwi.
CC      Pfam: PF02170; PAZ; 1.
CC      Pfam: PF02171; Piwi; 1.
CC      PROSITE: PS00821; PAZ; 1.
CC      PROSITE: PS00822; PIWI; 1.
CC      Initiation factor; Protein biosynthesis.
CC      Eukaryotic translation initiation factor
CC      CHAIN 1 860
CC      /FTId=PRO_0000194060.
CC
CC      DOMAIN 236 349
CC      FT DOMAIN 518 819
CC      FT PAZ.
CC      SQ SEQUENCE 860 AA; 97318 MW; A5B0798C66481C9C CRC64;
CC
Query Match 10.5%; Score 564; DB 1; Length 860;
Best Local Similarity 23.2%; Pred. No. 4.8e-23;
Matches 231; Conservative 156; Mismatches 358; Indels 250; Gaps 39;
QY      38 KKVLLLVNMFKFSKIYDEYEVYKMTKEVLNRKPGKPPFKTEIPIIDRAKLFQHL 97
Db      37 RTIKLQANFFEMDIPKIDIVHYELDIKPEK-----CPRVNEIVEHM-----V 80
QY      98 RHKKQTDFILEDY--VFDEKDTVYVCRINTVTSKMLVSEKVKVKDSEKKDEKLEKKI 155
Db      81 QHFKTK--IFGDRKEVFDGRKNLYTAMPLIGRDKV--ELEVTLPGEGRD--IFKVS 132
QY      156 LYTMLITRYKKPHLNFESRENPEKDEANRSYKELKVVMTQ--KRVYAPFVNEIKVQFAK 213
Db      133 IKVSCVSIQALHDAUGRLPSVPFE---TIQALDVVMRHLPSMRYTPV----- 178
QY      214 NFVYDNNISILRVPESEPHDNRFEQSLVAPRIEAFGIYIGIKELPDGEPVLNPAIVDKL 273
Db      179 -----GRSFTASEGCSNP-----LGGREVFEGFHQSVPSSL-WKWMNIDVSATA 224
QY      274 FYNAPKMSLLDYLILLVDPQSCNDVVDKDKTKLMAGKMTIROAARPIRQLLENLKLK- 332
Db      225 FYKA--QPVIEFVCEVLDFKSIIEEQ-----QKPLTDSQRVKFTKEIKGLKVEI 270
QY      333 --CAEWDN----ENSLRTERHLTFLDLCEENSLVYKVKTSKDRGNNAKYDTLTKIYE 386
Db      271 THCGQMKRYKRVNVTTRPASHQTF-PLQESGQTVECT-----VAQY----- 315
QY      387 ENKQFTEFPPLPLVKVSGAKAYAVPMHLEHVEHPQRYKNRIDLVMQDKFLKRAT-RKP 445
Db      316 RHKLVLRYPHLPCLQVQEQKHTYLPLEVCNIVAGQRCIKLTD--NQSTMTIRATARSA 373
QY      446 HDYKENTLKLKELDFSSSEELNFVERFGKSLQWIECPGKVLKPEMLVNSNEQIKMTPT 505
Db      374 PDQREISKLMSASFTNDP--YVREFGIMVKKDDMTDVTGRVLQPPSILYGRNKAIAATP 431
QY      506 V--IRGFQEKQLNV-----VPEKELCCAVFVNNET-----AGNP 537
Db      432 VQGVMDNRNKQFHTGIEIKVWALACFAPQRC-CTEVHLKSFTEQLKISRADGMPIQGQ 490

```

```
QY 538 CLEENDVVKFYTELIGCKFRGIRIGANENRGAQSI--MYDATKNEYAFYKKNCTLNTGIG 595
Db 491 CF-----CKYA-----QAGDSVEPMFRHLKNTYA-----514
QY 596 RPEIAATEAKNMFERLPDKEQKVLFIITIKRQLNAYGVKHYCHDHTIGVANQHSITSETV 655
Db 515 -----GLQLVVVLPGKTPVYAEVKRGVDTVLGMATQCVMKNV 553
QY 656 TKALASLRHEKSKRIFYQIALKINAKLGGINOELDWSIEAISPEEKERRKTMPLTMV 715
Db 554 -----QRTPTQTLNLCKINVLKGVNNIL-----LPQGRPPVFPQPV-IFL 595
QY 716 GIDVTHPTSYSGIDISIAAVASINPGGTIYRNMIIVTQECRPERGAHVAHGRERTDILEA 775
Db 596 GADVTHPPAGDGKKSIAAAGSMDAHNPYCATVRVQQ-----HREIITODLAA 645
QY 776 KPVKLLREPAENNDNRAPAHIVVYRDGSDSMLRVSHDELRLSKSEVKQFMSERDGEDP 835
Db 645 MYRELLIQYKTRFK-PRTIIFYRDGVSEGQFQVLHHELLAIREACIKLEKEY-----599
QY 836 EPKYTFIVIQKRHNTLLRLMEKDKPVPVKNDLTPAETDVAAVAKQWEDMKESKETGIV 895
Db 700 QPGITFIVQKRRHTRLF-CTDKNERVKGSGNIPA-----733
QY 896 NFSSGTTVDKLVSKYKPFPLASHGVLTGRPGHYTVMYDDKMGSDQDEVYKNTYGLAP 955
Db 734 -----GTTVDTKITHPTFPDFYLCSHAGIQGTSRPSHYHLWDDNRFSSDELQILTYQLCH 789
QY 956 LSARCKRPTSLVPVHYAHLSCAKELVRYTKEH 990
Db 790 TVYRCTRVSIIAPAYAHVAFRAR-YHLVDKEH 823

RESULT 15
Q4VAB3_MOUSE PRELIMINARY; PRT; 883 AA.
AC Q4VAB3;
DT 05-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Bif2c2 protein (Fragment).
GN Name=Bif2c2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B5/EGFP transgenic ICR mice; TISSUE=Trophoblast stem cells;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Alteschul S.P., Zesberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B5/EGFP transgenic ICR mice; TISSUE=Trophoblast stem cells;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC096465; AAH96465.1; -; mRNA.
DR SMR; QAVAB3; 246-413.
DR MGI; MGI:2446632; Bif2c2.
DR GO; GO:0003743; F:translation initiation factor activity; RCA.
DR GO; GO:0006412; P:protein biosynthesis; RCA.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Pwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Pwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; Pwi; 1.
DR NON_TER 1
SQ SEQUENCE 883 AA; 99333 MW; 6D807B35F5689017 CRC64;
Query Match 10.5%; Score 564; DB 2; Length 883;
Best Local Similarity 23.3%; Pred. No. 5e-23;
Matches 232; Conservative 157; Mismatches 356; Indels 250; Gaps 40;
QY 38 KVVLLVWVFESSKIYDREYEVKMTKEVLNRKPGKPPKTEIPIPDRAKLFWQHL 97
Db 60 RTIKLQANFFEMDIPKIDYHYELDI-----KPGK-CPRRVNREIVEHM-----V 103
QY 98 RHEKKQTDFILEDY--VFDEKQTVYVCRCLNTVTSKMLVSEKVKVKKDSEKDEKOLEKKI 155
Db 104 QHFKTQ---IFGDRKPVDFGRKNLYTAMPLPIGRDKV---ELEVLPGEKDR--IFKVS 155
QY 156 LYTMILTYRKPHLNFNSRENPEKBEANRSYKFLKNVMTQ--KVYAPVFNBEIKVQFAK 213
Db 156 IKWVSCVSLQALHDALSGRLSPVPE---TIQALDVMVHRHLPMSRYTPV-----201
QY 214 NFVYDNNILRVPSFHPDNPFEQSLVAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKL 273
Db 202 -----GRSFFTASEGCSNP-----LGGGREGVFGHQSVRPSL-WQOOLNIDVSATA 247
QY 274 FYNAPKMSLLDYLLIIVDPOSNDVVRKDLTKLMAGKMTIRQAARPRIRQLLENLKLK- 332
Db 248 FYKA--QPVEFVCEVLFKFSIEQ-----QKPLTDSQRVKFTKEIKGLKVEI 293
QY 333 --CAEVWDN----EMSLRTERHLTFDLCEENSLYKVTGKSDRGNAKKYDTTLFKIYE 386
Db 294 THCGQMKRKYRVCNVTRRPASHOTF-PLQOESGQTVECT-----VAQY---FK--D 338
QY 387 ENKKEIFEPHLPVLVKVKSAGKEYAVPMHELVHEKPYKRNIDLVMDQDKFLKRAT-RKP 445
Db 339 RHKLVLRYLPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLT--NOTSTWIRATARSA 396
QY 446 HDYKENTLMLKELDFSSSEELNFRVERFGLCSKLOMECPGKVLKBPMLVSNVNEQIKMTP 505
Db 397 PDRQBEISIKMKSASFNTDP--YVREFGIWKDEMTDVTGRVLQPPSILYGGNKAIATP 454
QY 506 V--IRGFOEKQLNV-----VPEKELCCAAVFFVNET-----AGNP 537
Db 455 VQGVWDMRNKQPHTGIEIKWAIACFAPQQR-CTEVHLKSFTEQLRKISRDAGMPTQGOP 513
QY 538 CLEENDVVKFYTELIGCKFRGIRIGANENRGAQSI--MYDATKNEYAFYKKNCTLNTGIG 595
Db 514 CF-----CKYA-----QAGDSVEPMFRHLKNTYA-----537
QY 596 RPEIAATEAKNMFERLPDKEQKVLFIITIKRQLNAYGVKHYCHDHTIGVANQHSITSETV 655
Db 538 -----GLQLVVVLPGKTPVYAEVKRGVDTVLGMATQCVMKNV 576
QY 656 TKALASLRHEKSKRIFYQIALKINAKLGGINOELDWSIEAISPEEKERRKTMPLTMV 715
Db 577 -----QRTPTQTLNLCKINVLKGVNNIL-----LPQGRPPVFPQPV-IFL 618
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:34:06 ; Search time 52 Seconds
(without alignments)
1716.947 Million cell updates/sec

Title: US-10-645-746-3
Perfect score: 5349
Sequence: 1 MSNPFPEKGGYRHSLDPE.....RHEMEHFLQTNVKYPMGSPA 1020

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB.pap.*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB.pap.*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB.pap.*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB.pap.*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/pCTUS COMB.pap.*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB.pap.*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	577	10.8	857	2	US-09-949-016-6819
2	577	10.8	894	2	US-09-949-016-10324
3	575.5	10.8	969	2	US-09-533-029-106
4	448.5	8.4	491	2	US-09-270-767-43512
5	427	8.0	1145	3	US-09-866-557A-5
6	391.5	7.3	623	2	US-09-270-767-43950
7	312	5.8	775	2	US-10-043-774B-2
8	306	5.7	861	2	US-09-873-737A-6
9	304	5.7	862	2	US-09-873-737A-4
10	288	5.4	829	2	US-10-104-047-3439
11	266.5	5.0	852	2	US-10-104-047-3646
12	265	5.0	257	2	US-09-248-796A-20922
13	209.5	3.9	844	2	US-10-043-774B-13
14	205	3.8	843	2	US-09-873-737A-2
15	178.5	3.3	245	2	US-09-270-767-44856
16	164	3.1	2954	3	US-09-150-867-1
17	163	3.1	2954	3	US-09-724-584-1
18	163	3.0	1780	2	US-09-949-016-6899
19	163	3.0	1786	2	US-09-949-016-7880
20	162.5	3.0	1244	2	US-09-349-016-11702
21	162	3.0	2662	2	US-09-595-684B-31
22	161.5	3.0	2663	2	US-09-538-092-1252
23	158	3.0	2733	2	US-09-949-016-11433
24	158	3.0	3259	2	US-09-949-016-6507
25	157	2.9	3248	1	US-08-353-700-1
26	157	2.9	3248	5	PCT-US95-16216-1

ALIGNMENTS

RESULT 1
US-09-949-016-6819
; Sequence 6819, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6819
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6819

Query Match 10.8%; Score 577; DB 2; Length 857;
Best Local Similarity 23.0%; Pred. No. 4.7e-43;
Matches 227; Conservative 171; Mismatches 365; Indels 222; Gaps 37;

QY	38	KKVLLVNMFKSSKIYDREYVEYKMTKEVLNRPKPGFPKKTIPIDRAKLFWQHL	97
DB	34	KPKILLANFEVDIPKIDVHYEVDIKPKD-----CPRVRREVVE-----YMW	77
QY	98	RHEKQTDFILEDY--VFDEKDTVSVCLNTVTSMVSEKV---VKDSEKKDKDLE	152
DB	78	QHFKQ---IFGDRKPVYDGNKIY-----TVPALPIGNERVDFVFTIPGEGKD-RIFK	127
QY	153	KKILYTMILTYRKCKFLNFSRNPKEKDBANRSYKFLKNVMTOKRYAPVFNBEIKVQFA	212
DB	128	VSIKMLAIVSWRLHEALVSGQIPVPLESQALDVAHRHLAS--MRYTPV-----	175
QY	213	KNFVYDNNILAVPESFHPDNFEQSLVAPRIEAMFGIYIGIKELFDGEPVLNFAIVDK	272
DB	176	-----GRSFFGPPEGYHP-----LGGGREGVWFGHQSVRPAM-WMOMLINIVSAT	220
QY	273	LFVNAPFMILLDYLLILLIVDPQCNDDVRDKLTKLMAGKMTIRQAARPIRQLLENKLK	332
DB	221	AFYKA--QPVIEMCEVLIRNIDEQPK-----PLTDSQVRVFTKEIKGLKVE	266

RESULT 3

US-09-533-029-106
; Sequence 106, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MEI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 106
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1149
US-09-533-029-106

Query Match 10.8%; Score 575.5; DB 2; Length 969;
Best Local Similarity 23.2%; Pred. No. 7.9e-43;
Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;
QY 38 KVVLLVNNPKFSKIYREYEVKMTKEVLRKPKPPKTEIPIDRAKLFQHL 97
DB 158 KVVVRAN--HFLVQVADRLYHYDVSINPEVSK----- 190
QY 98 RHEKQTDFLEDYVDFDEKDTVSVCKLNTVTVSKMLVSEKVKVQKSE--KDKDKLEKK 154
DB 191 -----TVNRNVKLLVKN---YKDSLGGKSPAYD-GRK 220
QY 155 ILVTM--ILTYRKHEHLPNSRENPEKDEBANRSYKFLKNV-----MTQKRYA 200
DB 221 SLTAGPLPDSKEFVNVLAERKADSGSGKDRPFKVAVKVNTSTDLYQLQOFLDKOREA 280
QY 201 PFVNEIKVQFAKNFVYDNNLSILRVPSFHDNRFQSL-----EVAPRIEAWFGI 251
DB 281 FY--DTIQVLVDVLRDKPSNDYVSGRSP-----FHTSLGKDARDGELGDGIEYWRGY 333
QY 252 YIGIKELFDGEVLPVFAVDKLFYNAPKMSLLDYLIVDPQSCNDVDRKDKTKKLMAGK 311
DB 334 FQSLR-LTQMLSLINDVSARSFY-----EPVVTDPISKFLNIRD--N 375
QY 312 MTRQARPRIROLLENKLKCAEVVDNEMSLRTERHLTLDLCEENSLVYKTKGSD-- 369
DB 376 RPLRDSRLKVKVRLTKVKKLH-WNGTKSA-----KISGISLSP 415
QY 370 -----RGRNAKCYDTTLFKIYEENKCF-IEPPHPLVVKVSGAKAYAVPMHEHLEHVEKQ 424
DB 416 IRELFTLEDKSEKTVVQVFAEKYNYRVKQALPAIQGTGSDTRPVLPMLCQIDE-QGR 474
QY 425 YKRIIDLVMQDKFLKATRKPHYKENTLQMLKELDFSSBELNFBVRFGKSLQMIETCP 484
DB 475 YTKRLNEKQVTKLLKATCORPPD-RENSIKMLVVQNNYNDLS--KEFGMSVTTQLASIE 531
QY 485 GKVLKEPMLVNSVNEQIKMTPTVIRGFEQKQLNVVPEKLCCEAVFVNVNETAGNCLSENDV 544

DB 532 ARVLPPMLKYTHDSGKEMVNPRLG-----QNNMIDKK----- 564
QY 545 VKFYTELIGCKFRGIRIGANENRGAQSIMYDATKNBYAFYKNCNTLNTGICRFEIAATEA 604
DB 565 -----NVNGAK---VTSWTCFEPQPAIPFISCPPEH-----IBEALLD- 600
QY 605 KNMFERLPDKQKVLMPFIISKROLNAYGFVKHYCDHTIGVANOHITSETVTYKALASLRH 664
DB 601 --IHKRAPGLQ---LLVILPVDVTSYGKIKRICETELGIVSQCCQPRQVKNL----- 648
QY 665 EKGSKRIFYOIALKINAKLGGINQELDWSHIAEISPEKERRKTMPL-----TWVYGDV 719
DB 649 ---NKQYMNVALKINVTGRTVLN-----DAIRRNIPLTITDPTIIMGADV 694
QY 720 THPTSYSGIDYSIAAVVASIN-PCGTTIYRNMTVQ---EE-----CRGERAVAHGR 767
DB 695 THPQGEDSSPSIAAVVASMDPEINKYRGVLSVAQAAREEIIQDLYKLQVDPQRLVH-- 752
QY 768 ERTDILEAKFVKLLRBEFAENNDNRAPAHIVVYRDGVSDESMRLRVSHDELRLSKSEVKQFM 827
DB 753 --SGLIREHFIAFRRA---TGQIPQRIIFYRDGVSSEGQPSQVLLHEMTAIRKACNSLQ 805
QY 828 SERDGEDPEPKYTFIVIOKEHNTLLRRMEKDKPVVKNDLTPAETDVAAVAKQWEEDMK 887
DB 806 -----ENYVPRVTFVIVQKHHTLRLFPBQHG-----NRDMT----- 836
QY 888 ESKETGIVNPSSGTTVDKLVSKYKFPFLASHHGLVGTSPGHYTYVMYDDKGMQSDRVY 947
DB 837 --DKSGNIQP--GTVDTKLCHNPEFDFYLNHAGIOGTSRPAHYHVLLENGGTAQOLQ 892
QY 948 KMTYGLAFLSARCKPISLPVHYAHLSCEKAKELYRTYKEHYHIGYVAQPRTR 1001
DB 893 MLTNNLCYTYARCTKSIVPPAYAHAAFPRA---RYTMESEMSDGGSSRSR 942

RESULT 4

US-09-270-767-43512
; Sequence 43512, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 43512
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43512

Query Match 8.4%; Score 448.5; DB 2; Length 491;
Best Local Similarity 27.5%; Pred. No. 9.4e-32;
Matches 148; Conservative 87; Mismatches 205; Indels 99; Gaps 16;
QY 467 NFVERFGLSKLQWIECPGKVLKPEMLVNSVNEQIKMTPTVIRGFEQKQLNVVPEKLCCEA 526
DB 8 SYVQEEGLTISNSMEVGRVLPKPKLOYGGR-----VSTGLTGQOL-FPPQNKVSLA 59
QY 527 VFVNVNAGNCPLEENDVVKFYTELIGGCKFRGIRIGANENRGAQ-SIMYDATKNYAFY 585
DB 60 -----SPNOGVMDMRGQKQFFTG-----BIRIWAJACFAPQRTVREDALRNFTQOL 105
QY 586 KNCITLNTG---IG-----RPEIAATEAKMPEBLPKQKVLMPFIISKQLNAYGFVKH 637
DB 106 QKTSNDAGMPIIGQPCFCFYATGPDQVEPMFRLYKTFPGLQLVWVWVLPKGTPTVYAEVKR 165
QY 638 YCDHTTIGVANOHITSETVTYKALASLRHESKGRIFVQIALKINAKLGGINQELDWSHIAE 697
DB 166 VGDTVLGMATQCQVQAKVNT-----SPOTLSNLCIKINVKLGGINSL----- 209

Qy	698	ISPEBKERRKMTPLTWYGVIDVTHPTSYSGIDYSAAVVASINPGGTIVRNMIVTQBECR	757
Db	210	-VPSTRPKVFNEPV-IFLGADVTHPAGDNKKPESIAAVVGSMDAHPRSRYAATVRVQQ----	264
Qy	758	PGERAVAHGRERTOLEAKFVKLLREFAENNDRAPAHITVVYRDGVSDSEMRLRVSHDEL	817
Db	265	-----HQYIIQELSSNVRELLIMFYKSTGGYKPHRIILLYRDGVSEGQPPHVLQHELT	317
Qy	818	SLKSEVKQFMSEGDGEDPE--PKYTFIVIQKHNTLLRRRMBKQKPVVNKOLTPAETDVA	875
Db	318	AIREACIKL-----EPEYRPGITFIVVQKRHHTLFCABEKE-----	354
Qy	876	VAAVKQWEEDMKESKETGIVNPSSGTTVDKILTVSKYKDPDFFLASHGVLTGSRPGHYTVM	935
Db	355	-----QSGKSG--NIPAGTTVDVGITHTEPDFYLCSHQGIQGTSRPSHYVL	400
Qy	936	YDDKMGSDQEVYKMTYGLAFLSARCKPISLPVPVHYAHLSEKAKELVRTYKHEYIGD	994
Db	401	WDDNHFDSDLOCLTYQLCHTYVRCRSVSIAPAPAYYAHVAFRAR--YHLVEKEHDSGE	458

RESULT 5

US-09-866-557A-5

; Sequence 5, Application US/09866557A
; Patent No. 7029844
; GENERAL INFORMATION:
; APPLICANT: Hammond, S.
; APPLICANT: Hannon, G.
; APPLICANT: Beach, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA INTERFERENCE
; FILE REFERENCE: GNCA-P02-007
; CURRENT APPLICATION NUMBER: US/09/866,557A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/189,739
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/243,097
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-866-557A-5

Query Match	8.0%;	Score 427;	DB 3;	Length 1145;
Best Local Similarity	21.1%;	Pred. No. 3.5e-29;		
Matches	210;	Conservative 152;	Mismatches 381;	Indels 250; Gaps 37;
Qy	59	Y E Y E V M T K E V L N R K P G P P K K T E I P I P D R A K L F W Q H L R H E K K Q T D F I L E D Y V D E K O T	118	
Db	364	Y H Y D V K I ----- M P E R P K F Y Q A F Q F R V D Q L G G A V L A Y D G K A S	403	
Qy	119	V Y S V C R - - L N T V T S K M L Y S E - - K V V K D S K S K O B K D - - L E K K I L Y T M I L T Y R K K F L N F	171	
Db	404	C Y S V D K L P L N S Q N E P V T D R N G T L R Y I I E K T G S T I D L S L T Y M - - - - -	452	
Qy	172	S R E N P E K D E A N R S Y K F L K N Y M T Q K V R Y A P F V N E E I K V Q F A K N F Y V D N N S I L R Y P E G F - -	229	
Db	453	--- N D R I F D K P M A Q C V E V V L A S P C - - - - - H N K A I R V G R S F F K	488	
Qy	230	- H D P N R E Q S L E V A P R I E A W F G I Y I G I K E L F D G E P V L N F A I V D K L F Y N A P K M S L L D V L L L	288	
Db	489	M S D P N N R H - - - E L D D G Y E A L V G L Y Q A F - M L G D R P P L N V D I S H K S F - - P I S M P M I E Y L - -	539	
Qy	289	I V D P Q S C N D D V R K D K T K L M A G K W T I R Q A A R P R I Q L L E N L K L K A E V W D N E M S R L T E R H	348	
Db	540	----- E R F S L K A K I - N N T N L D Y S R R F L P P L G I N V - - V T P P Q S - - - - -	577	
Qy	349	L T F I D L C E N S L V Y K V T G S D R G N A K Y D - - - - - T T L F K T Y E N K K P I E P P H L P L V K V K	403	
Db	578	----- F O S A P R V Y R N G L S R A P A S S E F F H D G K K V T I A S Y F H S R N Y P L K P P O L H C L N V G	631	

QY	404	SGAKEAYAVPMEHLEHVEHKPQRYKNRIDLVNQ-DKFLKQATRKPDHVDYKENTLKMKLKELDFS	462
DB	632	SSIKSIILLPIELCSIEB--GOALNRKDGATQVANMIKYATSTNVVRKRKIMNLLQYPOHN	689
QY	463	SEELNFRVERGLCSKLQWIECPGKVLKPEMLVNSNEQIKWTPIVRGFQEB---KOLNVV	518
DB	690	LDPT--ISRGIRIANDFIVSTRVLSPPQV---EYHSKRFTVVKVGSWRMDGKMLEPK	744
QY	519	PEKELCAVAFV-----VNETA-----GNCLBENDVVKFYTELIGCGKPRGIRIGANEN	567
DB	745	PKAHKCAVLYCDPRSGRKMYVTQLNDGRLNISQKAVNI--SLDSDTVYRPF--TDDE	799
QY	568	RGASIMYDATQNEYAFYKNCCTLTNGIRFEIAATAKXMPERLPDKEQKVLMPILISKR	627
DB	800	RSLDTIFADLKRSHD-----LAIVIPQF	824
QY	628	QLNAYGFVKYCHDTIGVANOHTSETVTVKALASLRHEKSGKRIFYQIALKINAKLGGIN	687
DB	825	RI-SYDTIKQKAELOHGILTCIQKQFTVER-----KCNQOTIGNILLKINSKLGIN	875
QY	688	QELDWSBIASISPEEKERRKTMPLTMVYGIDVTHPTSYSGIDYSIAAVVASINPGGTIYR	747
DB	876	HKI-----KDDPLPMMKNTMYIGADVTHPSPDQREIPSVVGVAAASHDPYGASYN	925
QY	748	NMIVTQECRPGERAVAGHERTDILEAKF---VKLLREPAENNDRAPAHIVVYRDGVS	804
DB	926	MQYRLQ-----RGALFEEDMFPSITLEHLRVVKEYR-NAYPDHIIYYRDGVS	971
QY	805	DSEMLRVSHDLRLSKSEVQFNSERDGEDPEPKYTFIVIQKHNTRLLRMEKDKPVVN	864
DB	972	DGQPKIKNBLRCIKQAC-----DKVCKPKICCVIVVKRHHTFFFP-----	1014
QY	865	KDLTPAETDVAAVAKOMESDMKESKETGIVNPSGGTTVDKLVSKYKFDFFFLASHHGVL	924
DB	1015	-----SGDVTTSKNFNNVDP--GTWVDRTIVHPNEMQPFMVSGQAIQ	1054
QY	925	GTSPRGHYTMVDDKMGSDQEVYKMTYGLAFLSARCRKPISLPVPVHYAHLSCSEKAKELY	984
DB	1055	GTAKPTRYNVIENGTNGIDIDLQOOLTYNLCHMPFRNRSVSPAPAYLAHLVAARGR-VY	1113
QY	985	RTYKEHYIG--DYAOPRTRHEMHEFLQTNVKY	1014
DB	1114	LTGTRFLDLKKEVAK---RTIVPEFMKKNPMY	1143
RESULT 6			
US-09-270-767-43950			
; Sequence 43950, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of <i>Drosophila melanogaster</i>			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; CURRENT FILING DATE: 1999-03-17			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn ver. 2.0			
; SEQ ID NO 43950			
; LENGTH: 623			
; TYPE: PRT			
; ORGANISM: <i>Drosophila melanogaster</i>			
; FEATURE:			
; OTHER INFORMATION: Xaa means any amino acid			
US-09-270-767-43950			
Query Match 7.3%; Score 391.5; DB 2; Length 623;			
Best Local Similarity 22.8%; Pred. No. 2.2e-26;			
Matches 156; Conservative 107; Mismatches 269; Indels 151; Gaps 24;			
QY	361	VYKVTGSDRGRNAKKYD-----TTLPKIVIEENKKFIEPPHLPVVKVSGAKAYAVPMEH	415
DB	61	VYRVNGLSRAPASSETPEHDKGKVTIASYPHSRNYPLKFPQLHCLNVGSSIKILLPIEL	120

Qy	416	LEVHEKPOYKRIIDLWQ--DKFLKRAKTKPHDYKENTLKMUKELDIFSSEELN-FVERFG	473
Db	121	CSIEE--GOALNRKDGATQVANNMIYAK--PHRRTCESARLTKCNTSTTWDPTISRFG	176
Qy	474	LCSKLOMTECPKVLKEPMLVNSVNEQIKMTPIVRGFQE---KQLNVPEKELCNVFPV	529
Db	177	IRIANDFIWVSTRVLNPPQV---EYHSKRFTWVKNGSWRMQMKLEPKPAHKHCAVLVC	233
Qy	530	-----VNSTAGNPCLEENDVVKFYFTELIGGCKFRGIRIGANENRCAQIMYDA	577
Db	234	DPRSGRKMYAQLND-FGNLIISQKAVNI---TLDSDDVYTRPF---TDDERSLDTIFADL	287
Qy	578	TKNEVAFYKNCNTLNTIGRFEIATAEAKMWERLPDKQKVLMFIIISKRQLNAPGVKH	637
Db	288	KRSQHD-----LAIVIIQFRI-SYDTTIQ	311
Qy	638	YCDHTIGVANOHITSETVKALASLRHBKSGKRIFYQIALKINAKLGGINOBLDWSIAE	697
Db	312	KABLQHGLITQCIQKQFVER-----KCNQOTIGNILLKINSKLNINHKI-----	356
Qy	698	ISPKEKRRKTMPLTMVIGIDVTHPTSVSGIDYSAAVVASINPGGTYRNMIVTQBECR	757
Db	357	---KDDPRLPMKXNTWIGADVTHPSPPQRBIPSVVGVAASHDPYGASYNMYRLQ----	409
Qy	758	PGERAVAHGREKTDILRAKF---VKLLREFANNDNRAPAHIVYVRDGVSGVSEMLRVSHD	814
Db	410	-----RGALSEIDMFITLHLRVYKEYR-NAYPDHIIYYRDGVSGQPPKIKNE	459
Qy	815	ELRSLKSEVKQFMSEKEDGPEPKYFTFVIOKRHNTRLRLRMEXDKPVVNKDLTPAETDV	874
Db	460	ELRCIKQAC-----DKVGCKPKICCVIVVKEHHTFRFP-----	492
Qy	875	AAVAAKQWEEPMKESKETGIWNPSSGTTVDKLIIVSKYKDFDFLASHGHGVLGTSRPHVTV	934
Db	493	-----SGVDTTNSKFNWVDP--GTVVDRTIVHPNEMQPFWMVSHQAIQGTAKPTRXNV	542
Qy	935	MYDDKMSQDVBVYKNTYGLAFLSARCKRPISLPVPVVAHLSECKAKELYTYKEHYIG-	993
Db	543	IENTGNLIDILLQQLTYNLCHMFPCNRSVSYPAPAYLAHLVAABGR-VYLTGTNRFLDL	601
Qy	994	--DYAQRTRRHEHFLQTNVQY	1014
Db	602	KCEYAK---RTIVPEFMKCKNPMY	621

```

RESULT 7
US-10-043-774B-2
; Sequence 2, Application US/10043774B
; Patent No. 6900017
; GENERAL INFORMATION:
; APPLICANT: University of Illinois at Chicago
; APPLICANT: Sharma, Arun
; APPLICANT: Hoffman, Ronald
; TITLE OF INVENTION: HUMAN HEMATOPOIETIC GROWTH REGULATORY GENE AND USES
; FILE REFERENCE: MBHB: CU08/PPA
; CURRENT APPLICATION NUMBER: US/10/043,774B
; CURRENT FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-043-774B-2

```

Qy	475	CSKLQMIQPCGKVLKEPMLVNSVNEQ-----IKMTPVIRGFOEKOLN----	511
Db	263	QYNQEIITD-----LQKPVIVSQPKRRGCGGTLPQAPMLIPELCYLTGLTDQRNDFNVM	317
Qy	517	-----VPEKEBLCCAVFVNVETAGN--PCLEENDVVK-----	546
Db	318	KDLAVHTRLTPEQR-----QREVGRLIDIYHKNDNVQRELRDGLSFDLSNLLSPSGR	369
Qy	547	-FYTELI--GCCKP--RGIRIGANENRGAOSIMYDATKNEVAFYKNCNTLTNGIGRPEJA	600
Db	370	ILQTEKHQGGKTFDYNQFADWSKETRGAPLISVKPLDNMLLIYTR-----RNYEAA	422
Qy	601	ATEAKNMPERLP-----DK-----EQKV-----LMPILISKEQLNAY	632
Db	423	NSLIQNLFKVTPAMQMRKAIMIEVDVDRTEAYLEVLQOKVTADTQIVVCLLSSNRKDKY	482
Qy	633	GFVKHYCDHTTGVANQHITSTVTYKALASLREHKSGKRIFYQIALKINAKLGGINQELDW	692
Db	483	DAIKKYLCTDCTPTPSQCVAARTLKG-----QQTWMAIATKIALQMNCKMGG---EL-W	531
Qy	693	SEIAEISPEEKERKTPW--LTMVVGIDVTHPTSYSGIDYSIAAVASINPGGTIYRNMI	750
Db	532	-----RVDPILKULMIVGIDCYDHMTAG--RRSTAGFVASINEGTMFSPRC	576
Qy	751	VTOBECRGERAVAHGRERTDILBAKFVKLIRFAENNDRNAPAHIVYVRDGVSDSEMLR	810
Db	577	IFQD-----RGQELVDGLKVCLOAALRANWSCNE--YMPSRILIVYRDGVGGQLXT	625
Qy	811	VSHDELRLSKSEVKQFMS--ERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPVVYNKDLT	868
Db	626	L-----VNYEVPQFLDCLKSIGRGYNPRLTVIWKRVNTRPE-----	663
Qy	869	PAETDVAAVAKQWEEDMKSEKTEG--IVNPSSCTTYVDKLIVSKYKPDFFLASHKGVLT	926
Db	664	-----AOSGGRQLQPLPGTVIDVEVTRPEWYDFFIVSQAVRSGS	702
Qy	927	SRPGHYTMVYDDKGMQSEVYKMTYGLAFLSARCKRKPISLPVPVHYAH	974
Db	703	VSPTHYNNVIYDNSGLKPDHQLRQTYKLCHYYNPGVIRVPAPCOYAH	750
RESULT 8			
US-09-873-737A-6			
; Sequence 6, Application US/09873737A			
; Patent No. 6723534			
; GENERAL INFORMATION:			
; APPLICANT: Duke University			
; APPLICANT: Lin, Haifan			
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE			
; TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SA			
; FILE REFERENCE: Artorney Docket No. 6723534 180-104/2			
; CURRENT APPLICATION NUMBER: US/09/873,737A			
; PRIOR FILING DATE: 2001-06-04			
; PRIOR APPLICATION NUMBER: PCT/US99/28764			
; PRIOR FILING DATE: 1999-12-03			
; PRIOR APPLICATION NUMBER: 60/110,901			
; PRIOR FILING DATE: 1998-12-04			
; NUMBER OF SEQ ID NOS: 21			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 6			
; LENGTH: 861			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (76)			
; OTHER INFORMATION: Xaa=Leu or Ile			
; NAME/KEY: misc_feature			
; LOCATION: (303)			
; OTHER INFORMATION: Xaa=Leu or Ile			
; NAME/KEY: misc_feature			
; LOCATION: (735)			
; OTHER INFORMATION: Xaa=Leu or Ile			

RESULT 8
 US-09-873-737A-6
 ; Sequence 6, Application US/09873737A
 ; Patent No. 6723534
 ; GENERAL INFORMATION:
 ; APPLICANT: Duke University
 ; APPLICANT: Lin, Haifan
 ; TITLE OF INVENTION: PURIFIED AND ISOLATED piw1 FAMILY GENES AND GENE
 ; TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
 ; FILE REFERENCE: Attorney Docket No. 6723534 180-104/2
 ; CURRENT APPLICATION NUMBER: US/09/873,737A
 ; CURRENT FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: PCT/US99/28764
 ; PRIOR FILING DATE: 1999-12-03
 ; PRIOR APPLICATION NUMBER: 60/110,901
 ; PRIOR FILING DATE: 1998-12-04
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 861
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (76)
 ; OTHER INFORMATION: Xaa=Leu or Ile
 ; NAME/KEY: misc_feature
 ; LOCATION: (303)
 ; OTHER INFORMATION: Xaa=Leu or Ile
 ; NAME/KEY: misc_feature
 ; LOCATION: (735)
 ; OTHER INFORMATION: Xaa=Leu or Ile

US-09-873-737A-6

Query Match 5.7%; Score 306; DB 2; Length 861;
Best Local Similarity 23.0%; Pred. No. 2.3e-18;
Matches 151; Conservative 88; Mismatches 226; Indels 174; Gaps 26;

QY 417 EVHE-KPORYKNRIDLVMDQKF-LKRAATKPHDYKENTLKLKELDFS-----462
DB 291 EHKFEQEQSVKEXIGLVLTKNYKTYRVDDIDWDONPKSTPKADGSGVSFLEYRKOY 350
QY 463 SEELNFERGLCSLQMIETCPKVLKEP-MLVNSVNEQIKMTPVIRGFOEQKQLNVVPEK 521
DB 351 NOBITDLKQVLVSQPKRRRGGTLPGPAMLIPELCYLTGLTDKMR-----NDFNMV--K 404
QY 522 ELCCAVFVNETAGNCP-----LEENDVVK-----FYTELI-- 552
DB 405 DLAVHTRLTPEQORQEVGRGLDIYIHKNDNVQRELDMGLSFDNSLLSFGSRILQTEKIHQ 464
QY 553 GGCKF---RGIRIGANENGAQSIMYDATKNEAFYKNCNTLNTGIRFPIAATAEAKNMP 609
DB 465 GGTFTDYNPQFADWSKETRGAPLISVKPLDNWLLIYTR-----RNYEAANSLLQNLFPK 517
QY 610 RLP-----DK-----EOKV-----LMFIIISKRQLNAYGFVKHYCDH 641
DB 518 VTPAMGMQMRKAIMIEVDRTTEAYLRVLQKVTADTQIVVCLLSNRKDKYDAIKKYLCT 577
QY 642 TIGVANQHITSETVTKALASLRHEKSGKEIFYQIALKINAKLGGINQELDMSEIAEISPE 701
DB 578 DCFPTSPQCVVARTLKG-----QOTVMAIATKIALQMNCKMGG--EL-W----- 617
QY 702 EKERRKTMF--LTMVYGVIDVTHPTSYSGIDYSIAAVVASINPGGTYIRNMIVTQBECPG 759
DB 618 ----RVDIPLKLVMIVGIDYCHDMTAG--RRSIAGFVASINEGTRFSCIFQD----- 666
QY 760 ERAVAHGRERTDILEAKFVKLLREFAEANDNRPAPAHIVYVRDGVSDSEMLRVSHDELRL 819
DB 667 ----RGQELVDGLKVCLOALRAWSGCNE-YMPSRIIVYRDGVSDGQDKTL-----V 713
QY 820 KSEVQKQFMS--ERDGEDPEPKYTFIVIOKRHNTRLLRRMEKDKPVVVKDLTPAETDVA 877
DB 714 NYEVQFQFLDCLKSIGRGYNPRUTVIVVKRVNARFF----- 749
QY 878 AVKQWEDMKESKETG--IVNPSGGTTVDKLIIVSKYKDFPFLASHHGVLGTSRPGHYTM 935
DB 750 -----AQSGRLQNLPGTVIDEVTREPWEYDFFIVSQAVRSGSVSPHYNVI 797
QY 936 YDCKGMSQDEVYKMTYGLAFLSARCKKPIISLPVPHYAH 974
DB 798 YDNSGLKPDHIQRLTYKLCIIYNNWPGVIRVPAPCOYAH 836

RESULT 9

US-09-873-737A-4
; Sequence 4, Application US/09873737A
; Patent No. 6723534
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE
; FILE REFERENCE: Attorney Docket No. 6723534 180-104/2
; CURRENT APPLICATION NUMBER: US/09/873,737A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110,901
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Mus sp.

FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (90)-
; OTHER INFORMATION: Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (216)
; OTHER INFORMATION: Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (383)
; OTHER INFORMATION: Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (816)
; OTHER INFORMATION: Xaa=Leu or Ile
US-09-873-737A-4

Query Match 5.7%; Score 304; DB 2; Length 862;

Best Local Similarity 22.0%; Pred. No. 3.5e-18;
Matches 141; Conservative 101; Mismatches 226; Indels 174; Gaps 24;

QY 414 EHLEVHE-KPORYKNRIDLVMDQKFLKRAATKPH-DYKENTLKLKELDFSSEELNFVER 471
DB 289 QOTEHKEQEVSKEGLIGLVLTKNYKTYRVDDIDWDONPKSTPKADGS--EVSFLEY 346
QY 472 FGLCSKLQMIETCPGKVLKEPMLVNSVNEQ-----IKMTPVIRGFOEQKQLN-- 516
DB 347 YRKQYNQBITD-----LKQPLVLSQPKRRRGGTLPGPAMXIPELCYLTGLTDKMRNDF 401
QY 517 -----VVPE---KELCCAVFVNETAGNCPLE-----ENDVVKFYTELI-- 552
DB 402 NVMKDLAVHTRLTPEQORQEVGRGLDIYTHKO--DNVQRELDMGLSFDNSLLSFGSRILOS 460
QY 553 ----GGCKF---RGIRIGANENGAQSIMYDATKNEAFYKNCNTLNTGIRFPIAATAEA 604
DB 461 EKIHOGGKTFDYNPQFADWSKETRGAPLISVKPLDNWLLIYTR-----RNYEAANSLLI 513
QY 605 KNMFRLP-----DK-----EOKV-----LMFIIISKRQLNAYGFVK 636
DB 514 QNLKPVTPAMGIQMKKAIMIEVDRTTEAYLRALQOKVTSDTQIVVCLLSNRKDKYDAIK 573
QY 637 HYCDHTIGVANQHITSETVTKALASLRHEKSGKEIFYQIALKINAKLGGINQELDMSEIA 696
DB 574 KYLTDCPTSPQCVVARTLKG-----QOTVMAIATKIALQMNCKMGGELWRVDMV-- 623
QY 697 EISPEKERRKTMPLTMVYGVIDVTHPTSYSGIDYSIAAVVASINPGGTYIRNMIVTQBERC 756
DB 624 -----LKLAMIVGIDYCHDHTTAG--RRSIAGFVASINEGTRFSCIFQD-- 667
QY 757 RPERAVAHGRERTDILEAKFVKLLREFAEANDNRPAPAHIVYVRDGVSDSEMLRVSHDEL 816
DB 668 -----RGQELVDGLKVCLOALRAWSGCNE-YMPSRVIYVRDGVSDGQDKTL----- 713
QY 817 RSLKSEVKQFMS--ERDGEDPEPKYTFIVIOKRHNTRLLRRMEKDKPVVVKDLTPAETDV 874
DB 714 --VNYEVQFQFLDCLKSIGRGYNPRUTVIVVKRVNARFF----- 750
QY 875 AVAAVKQWEDMKESKETG--IVNPSGGTTVDKLIIVSKYKDFPFLASHHGVLGTSRPGHY 932
DB 751 -----AQSGRLQNLPGTVIDEVTREPWEYDFFIVSQAVRSGSVSPHY 795
QY 933 TVMYDDKMSQDEVYKMTYGLAFLSARCKKPIISLPVPHYAH 974
DB 796 NVYDSSGLKPDHIQRLTYKLCIIYNNWPGVIRVPAPCOYAH 837

RESULT 10
US-10-104-047-3439
; Sequence 3439, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047

; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3439
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3439

Query Match 5.4%; Score 288; DB 2; Length 829;
Best Local Similarity 22.9%; Pred. No. 9.5e-17;
Matches 144; Conservative 87; Mismatches 207; Indels 192; Gaps 27;
Qy 417 EVHE-KPQRYKRRIDLVQDKFLKRAKPKH-DYKENTLKMELDFSSSEELNPFVFL 474
Db 291 EEHKFQVSKELIGLVLTAKYNNKTYRVDIDWDQNPSTFKADGS--EVSFLYRK 348
Qy 475 CSKLQMTCECPKVLKEPMLVNSVNEQ-----IKMTPVIRGFEKQLN----- 516
Db 349 QYNQSEITD-----LKQVLSQPKRRRPGGTLPGPAMLPELCYLTLGLTKMRNDFNM 403
Qy 517 -----VVPKEKLCVAVFVNAGN--PCLRENDVVK----- 546
Db 404 KDLAVHTRLTPEQR-----QREVGLTDYIHKNDNVQRELDRWGLSFDNLSFSGR 455
Qy 547 -FYTELI--GGCKP---RGIRIGANENRGAQSIMYDATKNEYAFYKNCNTLMTGIRFEIA 600
Db 456 ILQTEKIHQGGKTFDYNPQPADWSKETRGAPLISVKPLDNWLLIYR-----RNYEAA 508
Qy 601 ATEAKNMPERLP-----DK-----BQV-----LMFIIISKRLQNAV 632
Db 509 NSIQNLKFKVTPAMQMQRKAIMIEVDRTAEYLRVLQOKVTADTQIVWCLLSNRKDKY 568
Qy 633 GFVKHYCDHTIGVANQHTSETVTKALASLRHEKSKRIFOIALKINAKLGGINQELDW 692
Db 569 DAIKKYLCTDCTPSCQVARTLK-----QQTWAIATKIALQWCKNGG--EL-W 617
Qy 693 SEIAEISPEKERRKTMPP--LTMVYGDVTHPTSYSGIDYSIAAVASINPGGTIYENMI 750
Db 618 -----RVDIPLKLMIVGIDYHDMTAG--RESIAGFVASINEGTRWFSRC 662
Qy 751 VTQECRPGRAVAHGRERTDILEAKFVKLLRFAENNDNRAPAHIVVYRDGVSDEMLR 810
Db 663 IFQD-----RGQELVDGLKVCLOAALRAWNSCNE-YMPSRIITVYRDGVGDGLKT 711
Qy 811 VSHDELRLSKSVKQFMS--ERGEDPEPKVTFVIQKRNHNTLLRRMEKDKPVVWKDLT 868
Db 712 L-----VNYEVPQFLDCLKSIGRYNPRLTIVVKKRVNTRFF----- 749
Qy 869 PAETDVAAVAKQWEEDMKESKETG--IVNPSSGTTVDKLVSKYKFPDFFLASHHGVLT 926
Db 750 -----AQSGGRQLNPLGCTVIDVEVTRPEWYDFFIVSQVRSQS 788
Qy 927 SRPGHYTMDYDKMSQDEYVKTYYGLAFL 956
Db 789 VSPHYNVIYDNSGLKPDHIOQLTYKLCHI 818

RESULT 11
US-10-104-047-3446
; Sequence 3646, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: HL-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3646
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3646

Query Match 5.0%; Score 266.5; DB 2; Length 852;
Best Local Similarity 18.6%; Pred. No. 9.1e-15;
Matches 188; Conservative 140; Mismatches 316; Indels 369; Gaps 43;
Qy 40 VLLLVNMFKSSKIYDREYVEYKMTKEVLNRPKGPFPKKTIPDPRAKLFWQH--L 97
Db 106 VKLVNUNFNLDPP-QDWQLYQYHYVYIIPDLASRL-----RIALLYSHSEL 150
Qy 98 RHEKKQTFDILEDYVFDEKDTVSVCLNTVTSMVSEKVKVKDSEKDKLEKKILY 157
Db 151 SNEAKAFDGA-----LFLSQLEKVEKVELSSTQGETIKM 187
Qy 158 TMLTYRKFKHLPNSRENPKDBEANSYKFLKNVMTQKVRYAPFVNEEIKVQFAKNPVY 217
Db 188 TITL-----KRELPSSPVICQVF-----NIIPRK-----LKKLSMYQIGRNP-Y 227
Qy 218 DNNSILRVPSFHDPPRFEQSLVAPRIEAWFGIYIGIKELPDGEPVLNPAIVDKLFNA 277
Db 228 NPSEPMBIPOH-----KLSLWPGFATSV-SYFE-----RLKLFSA 261
Qy 278 PKMSLLDYLLLVDPQSCNDDVRKDKTKLKMAGQWTTIRQAARPRIROLLENLKLKCAEVM 337
Db 262 -----DVSYKVLNEVLEPWT-----ALCQRTGLSC----- 288
Qy 338 DNEMSLRTERHLFTLDCENSL-VYKVTGKSDRGNAKYDYTTLPKIYENKK----- 390
Db 289 -----FTQCEKQLIGLIVLTYYNNRTYSIDDDMSVKPHTTQKRGDGTET 335
Qy 391 FIBPPLPLVKVSKAKAYAVPMHEHLEVHEKPKQRYKNRIDLVMD-----KFLKRAKTRPH 446
Db 336 YVDY-----YKQYDITVSDLNQPMVLVSLKKRN 365
Qy 447 DYKENTL-KMLKELDFSSSEELNPFVERFGLCSKLQMIKCPGKVLKEPMLVNSVNEQIKWTP 505
Db 366 DNSEAQLAHLIPELCFLT-----GLTD-----QATSDFFQLMKVAEKTRLSLP 407
Qy 506 VIRGFEKQLNVPEKELCCAVFVNENAGNCPLENDVVKFYTELIGGCKFRIGIRGAN 565
Db 408 --SGRQORLARLDN-----IQRTNARFELET-----WGLHFGSQ 441
Qy 566 -----ENRGAQSI-MYDATKNEYAFYKNC-----TLNTGI-----GRPEI 599
Db 442 ISUTGRIVPSEKILMQDHICQPVSAADWSKD-----IRTCKILNAQSLNTWLILCSDRTEY 497
Qy 600 AATEAKNMPERL-----PDKEQKVLMPFIISKR 627
Db 498 VAESEFLNCLRRVAGSMGPNVDYPIIKVQENPAAFVRAIOQYVDPDQV---LWMCILPSN 554
Qy 628 QLNAYGFVKHYCHDTIGVANQHTSETVTKALASLRHEKSKRIFOIALKINAKLGGIN 687
Db 555 QKTYDYSIKKYLSSDCPVSPQCVLARTLNK-----QGMMSGIATKIAMQMTCKLGG-- 605
Qy 688 QELDWESEIAEISPEEKERRKTMPL--TMVYGDVTHPTSYSGIDYSIAAVASINPGGTI 745
Db 606 -EL-WA--VEI-----FLKSLMVVGVGDVCKDALS--DVMVVGCVASVNPRI 648
Qy 746 YRNVITQBECPRGERAVAHGRERTDILEAKFV---KLIRFEANNNDNRAPAHIVVYRD 801
Db 649 WFSRCILQ-----RTWTDVADCLKVFMTGALNKWYKYNHD--LPAHIIYRA 693
Qy 802 GVSDEMLRVSHDELRLSKSEVKQFMSERDGEDPEPKYTFIVIOKRNHTLLRRMEKDKP 861
Db 694 GVGDGQLETLIEYVFPQLLSVAE-----SGSNTSGRLSVIVVRKKCMRPFTEMTNT-- 746
Qy 862 VVWKDLTPAETDVAAVAKQWEEDMKESKETGIVNPSSGTTVDKLVSKYKFPDFFLASHH 921

Db 747 -----VQNPPLGTVDSEATRNWDYFYLISQV 774

QY 922 GVLGTSRPGHYTMVDDKMGSDQEVYKMYGLAFLSARCRKPISLDPVPHYAH 974

Db 775 ACRGTVSPTVYVNDNGLKPDHMQRLTFKLCHLYNNWPGIVSVDPAPQYAH 827

RESULT 12

US-09-248-796A-20922

; Sequence 20922, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 20922

; LENGTH: 257

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-20922

Query Match 5.0%; Score 265; DB 2; Length 257;

Best Local Similarity 26.0%; Pred. No. 1.7e-15;

Matches 75; Conservative 54; Mismatches 112; Indels 48; Gaps 9;

QY 714 YGIDVTHPTSGIDYSIAAVASI-NPGGTIYRNMIVTQECRPGERAVAHGRERTDI 772

Db 6 YLVLDVT--SSQGETYSVSSTIASIGSEDFGNKF-----PGSVRIQTGGQEV-I 54

QY 773 LEAKFVKLLR-EFAENNDNRPAHIVVYRDGVSDEMLRVSHDELRLSKSEVKQFMSERD 831

Db 55 ADVKSMVLERLENFHKIKGLPSKLVYRDGVSQGYTTILKEELTKIAAFNEYGLKLN 114

QY 832 GEDPEPKYFIVIQKRHNTRLLRMEKDKPVVKNKOLTPAETDVAVAVKQWBEEDMKESKE 891

Db 115 IPKYSPTIIFMIVVKRHTRFI-----PIHDN-----ADDPKTKQ 150

QY 892 TGIV---NPSSGTTVDKLVSKYKFFFLASHHGVLTGSRPGHYTMVDDKMGSDQEVYK 948

Db 151 IAVTSNENVIAGTVDREITSPAYDFYVQSQSLQGTGIPAHYVYLHDENNYTSDTIQK 210

QY 949 MTYGLAFLSARCRKPISLDPVPHYAHLSCEKAKELYRTYKEHYIGDYAQ 997

Db 211 ITYDLCHTFSRATKSVKVVPAAYADLLCTGRD-----YIYGFAK 251

RESULT 13

US-10-043-774B-13

; Sequence 13, Application US/10043774B

; Patent No. 6900017

; GENERAL INFORMATION:

; APPLICANT: University of Illinois at Chicago

; APPLICANT: Sharma, Arun

; APPLICANT: Hoffman, Ronald

; TITLE OF INVENTION: HUMAN HEMATOPOIETIC GROWTH REGULATORY GENE AND USES

; FILE REFERENCE: MBHB: CU08/PPA

; CURRENT APPLICATION NUMBER: US/10/043,774B

; CURRENT FILING DATE: 2002-06-06

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 13

; LENGTH: 844

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: PIWI protein

US-10-043-774B-13

Query Match 3.9%; Score 209.5; DB 2; Length 844;

Best Local Similarity 18.6%; Pred. No. 1.4e-09;

Matches 173; Conservative 140; Mismatches 283; Indels 335; Gaps 46;

QY 231 DPNRFEQSLVAPRIEA-----WFGIY-----IGIKELPDGEVVL---NFAIVD 271

Db 48 EASRERRALEEAPRREGGPPERKPGQDYDLNTRPVELSVKKGDTGVVLMQTQNFRLK 107

QY 272 KLFYNAPKMSLDYLLLVDPQSCDDVRKDKTK---LMAGKMTIRQAARPRIQLLE 327

Db 108 ----TKPEWRIVHYHVEP-EPSIENPRVVMGVLSNHNLLSGS-----YLPD 149

QY 328 NLKLCAEVWDMEMS-----RLTERHLTFLDLCEBNSL--VYKVTGKSDRGRN 373

Db 150 GLQLFTRKFEQEIIVLSGSKLDIEYKISIKFVGFISCAEPRFLQVLNLLRRSMKGLN 209

QY 374 AKKYDTTLFKIYEENKKTIEPPLPLVKVKSAAKEYAVPM-----EH-----LE 417

Db 210 LELVGRNLPD-----PRAKIE--IREFKMELWPGYETSIHQEKDILLGTE 253

QY 418 VHEKPORVKNRIDLVNQDKFLKRAKTRKPHDYKE---NTLKMKLKELDFSE--ELNFVER 471

Db 254 ITHKVMRTETIYDI-----MRRCSHNPARHQDEVRVNVLDLVLTDYNNRTYRINDVD- 306

QY 472 FGLCSKLOMIECPGKVL-----KEPMLVN----- 495

Db 307 FQOTPK-STFSCKGRDISVEYLYTKYNIRIRDHNPQLLISKNRDKALKTNASELVVLIP 365

QY 496 -----SVNEQIKMTPV-----IRGFOEKQKLVNVPKEKLCCEAVFV 530

Db 366 ELCRVTVGLNAEMRSFQLMRAMSSYTRMNPCKORTDLRAFNRLQN-TPE-----SVKVL 419

QY 531 NETAGNCPLEENDVVKFYTELIG--GCKPRGRIRIGANENRGAQSIWYDATKNEYAFYKNC 598

Db 420 RD--WNMELDKN-VTEVOGRIIGQONIVPHNCKVPAGENADWQRHFRDQR-----MLT 469

QY 599 TLNTGIGRFEIATE-----AKNMFERLPDKQKV----- 618

Db 470 TPSDGLDRWAVIAPQRNSHELRTLLDSLYRAASGMGLRIRSPQEFIIYDDRTGTYYVRAM 529

QY 619 -----LMFIIISKQLNAYFKVH--YCDHTIGVANOHITSETVTKALASLRHEKGS 668

Db 530 DCVRSDPKLILCLVPNDNAERYSSIKKRGYVDR--AVPTQVVTLKTKKPYSLM----- 591

QY 669 KRIFYQIALKINAKLGGINGNOELDWSEIAEISPEEKERRKTMPLT--MYVGIDVTHPTSYS 726

Db 582 -SIATKIALQNLCKLG-----YTPWMI-----LPLSGLMTIGFDIAKSTRDR 623

QY 727 GIDYSIAAVASINPGGTIYRN--MIVTOECRPGERAVAHGRERTDILE----AKFVKL 780

Db 624 KRAY--GALIASMD---LQONSTYFSTVTEC-----SAFDVLANTLWPMIAKA 666

QY 781 LREFAENNDNRPAHIVVYRDGVSDEMLRVSH---DELRSLSKSVKQFMSERDGEDPE 836

Db 667 LRQY-OHEHRKLPISRIVFYRDGVSGLSKQLFEFEVKDIEKLTAYARVQLS-----P 719

QY 837 PKYTFIVIOKRHNTRLLRMEKDKPVVKNKDLTPAETDVAVAVKQWBEEDMKESKETGINV 896

Db 720 FQLAIVVTRSNTRFNLNGQ----- 741

QY 897 PSSGTTVDKLVSKYKFPDFFLASHHGVLTGSRPGHYTMVDDKMGSDQEVYKMYGLAFL 956

Db 742 PPPGTIVDDVITLPERYDFYVVSQQVQGTVSPTSYNVLYSSMGLSPEKQKQLTKYKCHL 801

QY 957 ----SARCRKPISLDPVPHYAHLSCEKAKEL 983

Db 802 YNWSGTTTRVP-----AVCQYAKK. 821

Search completed: July 5, 2006, 13:35:41
Job time : 58 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 13:46:11 ; Search time 189 Seconds
(without alignments)
2499.890 Million cell updates/sec

Title: US-10-645-746-3

Perfect score: 5349

Sequence: 1 MSSNPFPELEKGFYRHSILDPE.....RHEMHPLOTNVVKYFGMSFA 1020

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /EMC_Celerra_SID33/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 2: /EMC_Celerra_SID33/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 3: /EMC_Celerra_SID33/ptodata/2/pubpaa/US09_PUBCOMB.pap:*
- 4: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10A_PUBCOMB.pap:*
- 5: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10B_PUBCOMB.pap:*
- 6: /EMC_Celerra_SID33/ptodata/2/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5349	100.0	1020	5	US-10-645-746-3
2	5349	100.0	1020	5	US-10-645-735-3
3	5349	100.0	1020	6	US-11-144-985-3
4	4271	79.8	818	5	US-10-645-746-13
5	4271	79.8	818	5	US-10-645-735-13
6	4271	79.8	818	6	US-11-144-985-13
7	575.5	10.8	969	3	US-09-533-029-106
8	575.5	10.8	969	4	US-10-286-264-102
9	575.5	10.8	969	4	US-10-374-780A-2084
10	575.5	10.8	969	4	US-10-412-699B-752
11	566	10.6	860	4	US-10-467-397-7
12	560.5	10.5	950	6	US-11-097-143-27930
13	560.5	10.5	984	6	US-11-097-143-14511
14	560.5	10.5	984	6	US-11-097-143-27927
15	557	10.4	766	5	US-10-645-746-9
16	557	10.4	766	5	US-10-645-735-9
17	557	10.4	766	6	US-11-144-985-9
18	555.5	10.4	737	5	US-10-645-746-10
19	555.5	10.4	737	5	US-10-645-735-10
20	555.5	10.4	737	6	US-11-144-985-10
21	555.5	10.4	860	5	US-10-756-149-4706
22	555	10.4	924	4	US-10-408-765A-995
23	549.5	10.3	816	5	US-10-483-505-31
24	536.5	10.0	1034	4	US-10-437-963-192340
25	536	10.0	949	4	US-10-437-963-108712
26	533.5	10.0	979	4	US-10-437-963-121457
27	519	9.7	1048	4	US-10-174-363-56

28	519	9.7	1048	4	US-10-374-780A-2086	Sequence 2086, Ap
29	519	9.7	1048	6	US-11-093-888-56	Sequence 56, Appl
30	509.5	9.5	990	4	US-10-437-963-108714	Sequence 108714,
31	507	9.5	1109	4	US-10-437-963-124320	Sequence 124320,
32	502	9.4	909	4	US-10-437-963-200613	Sequence 200613,
33	499.5	9.3	935	4	US-10-437-963-138118	Sequence 138118,
34	499.5	9.3	1088	4	US-10-437-963-158455	Sequence 158455,
35	496.5	9.3	1125	4	US-10-425-114-51644	Sequence 51644, A
36	495	9.3	577	4	US-10-174-363-2	Sequence 2, Appli
37	495	9.3	577	6	US-11-093-888-2	Sequence 2, Appli
38	490	9.2	1053	4	US-10-437-963-203412	Sequence 203412,
39	489	9.1	1101	4	US-10-174-363-54	Sequence 54, Appl
40	489	9.1	1101	6	US-11-093-888-54	Sequence 54, Appl
41	488	9.1	988	4	US-10-225-066A-152	Sequence 152, App
42	488	9.1	988	4	US-10-374-780A-238	Sequence 238, App
43	488	9.1	988	5	US-10-225-066A-152	Sequence 152, App
44	487	9.1	821	4	US-10-174-363-4	Sequence 4, Appli
45	487	9.1	821	6	US-11-093-888-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-645-746-3
; Sequence 3, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabata, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; FILE REFERENCE: UMY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-746-3

Query Match 100.0%; Score 5349; DB 5; Length 1020;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MSSNPFPELEKGFYRHSILDPEMKWLARPTGCDKGKFKVLLVNVNFKFSKIYDREYYE	60		
Db	1	MSSNPFPELEKGFYRHSILDPEMKWLARPTGCDKGKFKVLLVNVNFKFSKIYDREYYE	60		
Qy	61	YEVMTKEVLNKRPGPKKTRIPIDRAKLFQHLRHEKKOTDFTLEDVVPDEKDTVY	120		
Db	61	YEVMTKEVLNKRPGPKKTRIPIDRAKLFQHLRHEKKOTDFTLEDVVPDEKDTVY	120		
Qy	121	SVCLRLNTVTSKMLVSEKVKVSKDEKDEKLEKKILYTLTYRKKFHLNFSRENPEKDE	180		
Db	121	SVCLRLNTVTSKMLVSEKVKVSKDEKDEKLEKKILYTLTYRKKFHLNFSRENPEKDE	180		
Qy	181	EANRSYKFLKNVMTQKRVAPFNVEIKVQFAKNFYVNNNSILRVPSFHDNRFESGLE	240		
Db	181	EANRSYKFLKNVMTQKRVAPFNVEIKVQFAKNFYVNNNSILRVPSFHDNRFESGLE	240		
Qy	241	VAPRIEAWFGIYIGIKELPDGEPLNFAIVDKLFYNAPKMSLADYLLLIYDQSCNDVVR	300		
Db	241	VAPRIEAWFGIYIGIKELPDGEPLNFAIVDKLFYNAPKMSLADYLLLIYDQSCNDVVR	300		

```

Db 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLLIIVDPQSCNDDVR 300
Qy 301 KDLTKTLMAGKMTIRQAARPRIRQLLENLKLKCAEVWDMESRLTERHLTFLDLCENSL 360
Db 301 KDLTKTLMAGKMTIRQAARPRIRQLLENLKLKCAEVWDMESRLTERHLTFLDLCENSL 360
Qy 361 VYKVTGKSDRGNAKKYDTTLFKIYBENKCFTEFPHLPLVKVKGAKAYAVPMHLEVHE 420
Db 361 VYKVTGKSDRGNAKKYDTTLFKIYBENKCFTEFPHLPLVKVKGAKAYAVPMHLEVHE 420
Qy 421 KPORYKNRIDLVMQDKFLKRAKTRKPHDYKENTLKMELDFSSSEELNFVERFGLCSKLM 480
Db 421 KPORYKNRIDLVMQDKFLKRAKTRKPHDYKENTLKMELDFSSSEELNFVERFGLCSKLM 480
Qy 481 IECPGKVLKEPMLVNSVNEIQIKNMTPIRGFQEKQLNVVPEKELCCAVFVNETAGNPCLE 540
Db 481 IECPGKVLKEPMLVNSVNEIQIKNMTPIRGFQEKQLNVVPEKELCCAVFVNETAGNPCLE 540
Qy 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Db 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Qy 601 ATEAKMFERLPDKEQKVMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALA 660
Db 601 ATEAKMFERLPDKEQKVMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALA 660
Qy 661 SURHEKSGRIFYQIATALKINAKLGGINQELDSEIAEISPEKERRKTMPLTMYVGIDVT 720
Db 661 SURHEKSGRIFYQIATALKINAKLGGINQELDSEIAEISPEKERRKTMPLTMYVGIDVT 720
Qy 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQECPGERAVAHGRETRDILEAKFVKL 780
Db 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQECPGERAVAHGRETRDILEAKFVKL 780
Qy 781 LREFAENNDNRAPAHIVVYRDGVSDEMLRVSHDELRLSKSVKQFMSRDEGEDPEPKYT 840
Db 781 LREFAENNDNRAPAHIVVYRDGVSDEMLRVSHDELRLSKSVKQFMSRDEGEDPEPKYT 840
Qy 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEDMKESKETGI VNPSSG 900
Db 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEDMKESKETGI VNPSSG 900
Qy 901 TTVDKLIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMQDEYVYKMTYGLAFLSARC 960
Db 901 TTVDKLIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMQDEYVYKMTYGLAFLSARC 960

RESULT 2
US-10-645-735-3
; Sequence 3, Application US/10645735
; Publication No. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMG-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-735-3

Query Match
Best Local Similarity 100.0%; Score 5349; DB 5; Length 1020;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSNPPELEKGFYRHSLDPEMKWMLARPTGKCDGKFYEKKVLLLVNMFKSSKIYDREYVE 60
Db 1 MSSNPPELEKGFYRHSLDPEMKWMLARPTGKCDGKFYEKKVLLLVNMFKSSKIYDREYVE 60
Qy 61 YEVMKTEVLRNPKPKPPKTEIIPDRAKLFWOHLRHEKKOTDPILEDYVDEKDTVY 120
Db 61 YEVMKTEVLRNPKPKPPKTEIIPDRAKLFWOHLRHEKKOTDPILEDYVDEKDTVY 120
Qy 121 SVCRLNTVTSKMLVSEKVVVKDSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
Db 121 SVCRLNTVTSKMLVSEKVVVKDSEKKDEKLEKKILYTMILTYRKKFHLNFSRENPEKDE 180
Qy 181 EANRSYKFLKNVMTQKVYAPFVNBEIKVQFAKNFYDNNSILRVPESEPHDPNRFQESLE 240
Db 181 EANRSYKFLKNVMTQKVYAPFVNBEIKVQFAKNFYDNNSILRVPESEPHDPNRFQESLE 240
Qy 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLLIIVDPQSCNDDVR 300
Db 241 VAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYNAPKMSLLDYLLLIIVDPQSCNDDVR 300
Qy 301 KDLTKTLMAGKMTIRQAARPRIRQLLENLKLKCAEVWDMESRLTERHLTFLDLCENSL 360
Db 301 KDLTKTLMAGKMTIRQAARPRIRQLLENLKLKCAEVWDMESRLTERHLTFLDLCENSL 360
Qy 361 VYKVTGSDRGNAKKYDTTLFKIYBENKCFTEFPHLPLVKVKGAKAYAVPMHLEVHE 420
Db 361 VYKVTGSDRGNAKKYDTTLFKIYBENKCFTEFPHLPLVKVKGAKAYAVPMHLEVHE 420
Qy 421 KPORYKNRIDLVMQDKFLKRAKTRKPHDYKENTLKMELDFSSSEELNFVERFGLCSKLM 480
Db 421 KPORYKNRIDLVMQDKFLKRAKTRKPHDYKENTLKMELDFSSSEELNFVERFGLCSKLM 480
Qy 481 IECPGKVLKEPMLVNSVNEIQIKNMTPIRGFQEKQLNVVPEKELCCAVFVNETAGNPCLE 540
Db 481 IECPGKVLKEPMLVNSVNEIQIKNMTPIRGFQEKQLNVVPEKELCCAVFVNETAGNPCLE 540
Qy 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Db 541 ENDVVKFYTELIGGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCITLNTGIGRFEIA 600
Qy 601 ATEAKMFERLPDKEQKVMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALA 660
Db 601 ATEAKMFERLPDKEQKVMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALA 660
Qy 661 SURHEKSGRIFYQIATALKINAKLGGINQELDSEIAEISPEKERRKTMPLTMYVGIDVT 720
Db 661 SURHEKSGRIFYQIATALKINAKLGGINQELDSEIAEISPEKERRKTMPLTMYVGIDVT 720
Qy 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQECPGERAVAHGRETRDILEAKFVKL 780
Db 721 HPTSYSGIDYSIAAVASINPGGTIYRNMIVTQECPGERAVAHGRETRDILEAKFVKL 780
Qy 781 LREFAENNDNRAPAHIVVYRDGVSDEMLRVSHDELRLSKSVKQFMSRDEGEDPEPKYT 840
Db 781 LREFAENNDNRAPAHIVVYRDGVSDEMLRVSHDELRLSKSVKQFMSRDEGEDPEPKYT 840
Qy 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEDMKESKETGI VNPSSG 900
Db 841 FVIOQRHNTRLRRMEKDKPVVVKDLTPAETDVAVAVKQWEDMKESKETGI VNPSSG 900
Qy 901 TTVDKLIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMQDEYVYKMTYGLAFLSARC 960
Db 901 TTVDKLIVSKYKFDFFLASHHGVLGTSRPGHYTVMYDDKGMQDEYVYKMTYGLAFLSARC 960

```

Qy	961	RKPTSLPVPVHYAHLSCFKAKELRYTYKEHYIGDYAOPRTRHNEHMFLOTNVKYPGMSFA	1020
Db	961	RKPTSLPVPVHYAHLSCFKAKELRYTYKEHYIGDYAOPRTRHNEHMFLOTNVKYPGMSFA	1020

```

RESULT 3
US-11-144-985-3
; Sequence 3, Application US/11144985
; Publication No. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: RNA INTERFERENCE
; FILE REFERENCE: UMY-052CN
; CURRENT APPLICATION NUMBER: US/11/144,985
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-144-985-3

```

Query Match	100.0%;	Score 5349;	DB 6;	Length 1020;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1020;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSGNFPLEKGFYRHSJDPEMKMLARPTGCDGKFYEKKVLLLVNVPKSSKIYDREY	60		
Db	1	MSGNFPLEKGFYRHSJDPEMKMLARPTGCDGKFYEKKVLLLVNVPKSSKIYDREY	60		
Qy	61	YEVQMTKEVNLNRKPGKFPFKTPIPIPDRAKLFWQHLRHEKKQTDFILEDYVDEKDTVY	120		
Db	61	YEVQMTKEVNLNRKPGKFPFKTPIPIPDRAKLFWQHLRHEKKQTDFILEDYVDEKDTVY	120		
Qy	121	SVCELNTVTSKMLVSEKVKVKQDSKKDEKLEKKILYTMILTYRKPFHMLNFSRENPEKDE	180		
Db	121	SVCELNTVTSKMLVSEKVKVKQDSKKDEKLEKKILYTMILTYRKPFHMLNFSRENPEKDE	180		
Qy	181	EANRSYKFLKNVMTQKVRYAPFVNEEIKVQFANKFVYDNNNSILRVPESHDPNRPQSLE	240		
Db	181	EANRSYKFLKNVMTQKVRYAPFVNEEIKVQFANKFVYDNNNSILRVPESHDPNRPQSLE	240		
Qy	241	VAPRIEAWFGIYIGIKELFDGEPVLNPAIVDKLFYNAPKMSLLDYLLILVDPQSCNDDVR	300		
Db	241	VAPRIEAWFGIYIGIKELFDGEPVLNPAIVDKLFYNAPKMSLLDYLLILVDPQSCNDDVR	300		
Qy	301	KDLKTKLMAGKMTIROAARPRIRQLLENLKLCAEVMDNEMSKLTERHILTFLDLCBENS	360		
Db	301	KDLKTKLMAGKMTIROAARPRIRQLLENLKLCAEVMDNEMSKLTERHILTFLDLCBENS	360		
Qy	361	VYKVTGKSDRGRNAKYDITLTKIYENKXKFIFFPHLPVVKVSGKAEYAVPMEHLEVHE	420		
Db	361	VYKVTGKSDRGRNAKYDITLTKIYENKXKFIFFPHLPVVKVSGKAEYAVPMEHLEVHE	420		
Qy	421	KPQRYKNRIDLVMQDKFLKEATRKPXDHYKENTLKMJLKELDFFSSEELNFWERFGLCSKQ	480		
Db	421	KPQRYKNRIDLVMQDKFLKEATRKPXDHYKENTLKMJLKELDFFSSEELNFWERFGLCSKQ	480		
Qy	481	IECPGKVLKEPMLVNSVNEQIKMTPTVIRGQEKOLNVVPEKLCACAVFVNVTAGNPCI	540		
Db	481	IECPGKVLKEPMLVNSVNEQIKMTPTVIRGQEKOLNVVPEKLCACAVFVNVTAGNPCI	540		

Qy	541	ENDVVKFYTELIGGCKPGRIGIRGANENRGAQSIWYDATKNEYAFYKNCITLNTGIRFEIA	600
Db	541	ENDVVKFYTELIGGCKPGRIGIRGANENRGAQSIWYDATKNEYAFYKNCITLNTGIRFEIA	600
Qy	601	ATEAKNWERLPDKQKVLMPFIILSKQOLNAYGFVKHYCDHTTIGVANOHIITSETVTKALA	660
Db	601	ATEAKNWERLPDKQKVLMPFIILSKQOLNAYGFVKHYCDHTTIGVANOHIITSETVTKALA	660
Qy	661	SLRHEKSGKRIFYQIALKINAKLGGINGNOELDWESEIAEISPEEKERRKTMELTWYVGIDVT	720
Db	661	SLRHEKSGKRIFYQIALKINAKLGGINGNOELDWESEIAEISPEEKERRKTMELTWYVGIDVT	720
Qy	721	HPTSYSGIDYSIAAVASINPGGTIYRNMTVTOEECECPGERAVAHGERTDIILEAKFVKL	780
Db	721	HPTSYSGIDYSIAAVASINPGGTIYRNMTVTOEECPGERAVAHGERTDIILEAKFVKL	780
Qy	781	LREPAENNDNRPAHIWYVDGVSDESMLRVSHDELRSLSKSEVKQFMSERDGEDDPEPKYT	840
Db	781	LREPAENNDNRPAHIWYVDGVSDESMLRVSHDELRSLSKSEVKQFMSERDGEDDPEPKYT	840
Qy	841	FIVIQKHNRLLRMEKOKPVPVYNKOLTPAETDVAVAAVKQWBEEDMKESKETGIVNPPSSG	900
Db	841	FIVIQKHNRLLRMEKOKPVPVYNKOLTPAETDVAVAAVKQWBEEDMKESKETGIVNPPSSG	900
Qy	901	TTVDKLIYSKYKPDFFLASHHGVLTGTSRPGHYTMYDSDKGMSODEVYKMTYGLAPLSARC	960
Db	901	TTVDKLIYSKYKPDFFLASHHGVLTGTSRPGHYTMYDSDKGMSODEVYKMTYGLAPLSARC	960
Qy	961	RKPIISLPVPVHYAHLSCBAKELIYRTYKHYIGDYAOPRTRHEMEHFLOTNVKYPGMSFA	1020
Db	961	RKPIISLPVPVHYAHLSCBAKELIYRTYKHYIGDYAOPRTRHEMEHFLOTNVKYPGMSFA	1020

```

RESULT 4
US-10-645-746-13
; Sequence 13, Application US/10645746
; Publication No. US20040265839A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATH
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMY-052DV1
; CURRENT APPLICATION NUMBER: US/10/645,746
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-746-13

```

```

Query Match      79.8%; Score 4271; DB 5; Length 818;
Best Local Similarity 100.0%; Pred. No. 4.7e-316;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 203 VNEEIKVQFAKNFVYDNNISILRVPSFHDNPRFQSLEVPRIEAWFGYIYGIGELFDGE 262
Db 1 VNEEIKVQFAKNFVYDNNISILRVPSFHDNPRFQSLEVPRIEAWFGYIYGIGELFDGE 60

Qy 263 PVLNFAIVDKLFYNAPQMSLLDYLLLIIVDPQSCNDDVKDKLTKYLMAGQWTTIROAARPRI 322
Db 61 PVLNFAIVDKLFYNAPQMSLLDYLLLIIVDPQSCNDDVVKDLTKYLMAGQWTTIROAARPRI 120

```

```
QY 323 ROLLENKLCACAEVNDNMSRLTERHLTFDLDCSENSLVYKVTGKSDRGNAKKYDTTLF 382
DB 121 ROLLENKLCACAEVNDNMSRLTERHLTFDLDCSENSLVYKVTGKSDRGNAKKYDTTLF 180
QY 383 KIYEENKKFIEPPLPLVVKVSGAKAYAVPMEHLEVHEKPORYKNRIDLVMDQKFLKRAT 442
DB 181 KIYEENKKFIEPPLPLVVKVSGAKAYAVPMEHLEVHEKPORYKNRIDLVMDQKFLKRAT 240
QY 443 RPHDYKENTLKMKELDSSSEELNVERFGLCSKLQMIETCPGVKLEPMLVNSVNEQIK 502
DB 241 RPHDYKENTLKMKELDSSSEELNVERFGLCSKLQMIETCPGVKLEPMLVNSVNEQIK 300
QY 503 MTPVIRGFOEKQNLNVPEKELCCAVFVNAGNCPCLBENDVVKFTELIIGCKRGIRI 562
DB 301 MTPVIRGFOEKQNLNVPEKELCCAVFVNAGNCPCLBENDVVKFTELIIGCKRGIRI 360
QY 563 GANENRGAQSI MYDATKNEYAFYKNCITLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 622
DB 361 GANENRGAQSI MYDATKNEYAFYKNCITLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 420
QY 623 IISKQNLNAYGFVKHYCDHTIGVANQHIITSETVTKALASLRHEKSGSKRIFYQIALKINAK 682
DB 421 IISKQNLNAYGFVKHYCDHTIGVANQHIITSETVTKALASLRHEKSGSKRIFYQIALKINAK 480
QY 683 LGGINQELDWSIEAISIPEEKERRKTMPLTMVVGIDVTHPTSYSGIDYSIAAVASINPG 742
DB 481 LGGINQELDWSIEAISIPEEKERRKTMPLTMVVGIDVTHPTSYSGIDYSIAAVASINPG 540
QY 743 GTIYRNMIVTQBECPGERAVAHGRERTDILEAKFVKLLRFEAENNDNRAPAHIVVYRDG 802
DB 541 GTIYRNMIVTQBECPGERAVAHGRERTDILEAKFVKLLRFEAENNDNRAPAHIVVYRDG 600
QY 803 VSDSEMLRVSHDELRSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 862
DB 601 VSDSEMLRVSHDELRSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 660
QY 863 VNKDLTPAETDVAAVAKQWEEDMKESKETGIVNPSGGTTVDKLI VSKYKFPDFFLASHHG 922
DB 661 VNKDLTPAETDVAAVAKQWEEDMKESKETGIVNPSGGTTVDKLI VSKYKFPDFFLASHHG 720
QY 923 VLGTSRPGHYTVMYDDKQMSQDEVYKMTYGLAFLSARCKRPISLPVPVHYAHLSCAKE 982
DB 721 VLGTSRPGHYTVMYDDKQMSQDEVYKMTYGLAFLSARCKRPISLPVPVHYAHLSCAKE 780
QY 983 LYRTYKHYIGDYAOPRTRHEMEHFLQTNVVKPGMSFA 1020
DB 781 LYRTYKHYIGDYAOPRTRHEMEHFLQTNVVKPGMSFA 818
```

RESULT 5

```
US-10-645-735-13
; Sequence 13, Application US/10645735
; Publication NO. US20050100913A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMG-052
; CURRENT APPLICATION NUMBER: US/10/645,735
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/689,992A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
```

; LENGTH: 818

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-10-645-735-13

Query Match

Best Local Similarity 79.8%; Score 4271; DB 5; Length 818;

Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 VNEEIKVQFAKNFVYDNNNSILRVPSFHDNPRFEQSLVAPRIEAWFGIYIGIKELFDGE 262

DB 1 VNEEIKVQFAKNFVYDNNNSILRVPSFHDNPRFEQSLVAPRIEAWFGIYIGIKELFDGE 60

QY 263 PVLNFAIVDKLPYNAPKMSLLDYLLIIVDPQSCNDDVRKDLTKLMAGQWTTIROAAPRI 322

DB 61 PVLNFAIVDKLPYNAPKMSLLDYLLIIVDPQSCNDDVRKDLTKLMAGQWTTIROAAPRI 120

QY 323 ROLLENKLCACAEVNDNMSRLTERHLTFDLDCSENSLVYKVTGKSDRGNAKKYDTTLF 382

DB 121 ROLLENKLCACAEVNDNMSRLTERHLTFDLDCSENSLVYKVTGKSDRGNAKKYDTTLF 180

QY 383 KIYEENKKFIEPPLPLVVKVSGAKAYAVPMEHLEVHEKPORYKNRIDLVMDQKFLKRAT 442

DB 181 KIYEENKKFIEPPLPLVVKVSGAKAYAVPMEHLEVHEKPORYKNRIDLVMDQKFLKRAT 240

QY 443 RPHDYKENTLKMKELDSSSEELNVERFGLCSKLQMIETCPGVKLEPMLVNSVNEQIK 502

DB 241 RPHDYKENTLKMKELDSSSEELNVERFGLCSKLQMIETCPGVKLEPMLVNSVNEQIK 300

QY 503 MTPVIRGFOEKQNLNVPEKELCCAVFVNAGNCPCLBENDVVKFTELIIGCKRGIRI 562

DB 301 MTPVIRGFOEKQNLNVPEKELCCAVFVNAGNCPCLBENDVVKFTELIIGCKRGIRI 360

QY 563 GANENRGAQSI MYDATKNEYAFYKNCITLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 622

DB 361 GANENRGAQSI MYDATKNEYAFYKNCITLNTGIGRFEIAATEAKNMPERLPDKEQKVLMEI 420

QY 623 IISKQNLNAYGFVKHYCDHTIGVANQHIITSETVTKALASLRHEKSGSKRIFYQIALKINAK 682

DB 421 IISKQNLNAYGFVKHYCDHTIGVANQHIITSETVTKALASLRHEKSGSKRIFYQIALKINAK 480

QY 683 LGGINQELDWSIEAISIPEEKERRKTMPLTMVVGIDVTHPTSYSGIDYSIAAVASINPG 742

DB 481 LGGINQELDWSIEAISIPEEKERRKTMPLTMVVGIDVTHPTSYSGIDYSIAAVASINPG 540

QY 743 GTIYRNMIVTQBECPGERAVAHGRERTDILEAKFVKLLRFEAENNDNRAPAHIVVYRDG 802

DB 541 GTIYRNMIVTQBECPGERAVAHGRERTDILEAKFVKLLRFEAENNDNRAPAHIVVYRDG 600

QY 803 VSDSEMLRVSHDELRSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 862

DB 601 VSDSEMLRVSHDELRSKSEVKQFMSERDGEDPEPKYTFIVIQKRHNTLLRRMEKDKPV 660

QY 863 VNKDLTPAETDVAAVAKQWEEDMKESKETGIVNPSGGTTVDKLI VSKYKFPDFFLASHHG 922

DB 661 VNKDLTPAETDVAAVAKQWEEDMKESKETGIVNPSGGTTVDKLI VSKYKFPDFFLASHHG 720

QY 923 VLGTSRPGHYTVMYDDKQMSQDEVYKMTYGLAFLSARCKRPISLPVPVHYAHLSCAKE 982

DB 721 VLGTSRPGHYTVMYDDKQMSQDEVYKMTYGLAFLSARCKRPISLPVPVHYAHLSCAKE 780

QY 983 LYRTYKHYIGDYAOPRTRHEMEHFLQTNVVKPGMSFA 1020

DB 781 LYRTYKHYIGDYAOPRTRHEMEHFLQTNVVKPGMSFA 818

RESULT 6

```
US-11-144-985-13
; Sequence 13, Application US/11144985
; Publication NO. US20060024798A1
; GENERAL INFORMATION:
; APPLICANT: Mello, Craig C.
; APPLICANT: Tabara, Hiroaki
```

; APPLICANT: Grishok, Alla
; APPLICANT: Fire, Andrew
; TITLE OF INVENTION: RNA INTERFERENCE PATHWAY GENES AS TOOLS FOR TARGETED GENETIC
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: UMY-052CN
; CURRENT APPLICATION NUMBER: US/11/144,985
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 09/689,992
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/193,218
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/159,776
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-144-985-13

Query Match 79.8%; Score 4271; DB 6; Length 818;
Best Local Similarity 100.0%; Pred. No. 4.7e-316;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 203 VNEEIKVQAKNFYDNNILRVDSFHDNRFQSLVAPRIEAWFGIYIGIKELFDGE 262
Db 1 VNEEIKVQAKNFYDNNILRVDSFHDNRFQSLVAPRIEAWFGIYIGIKELFDGE 60

Qy 263 PVLNFAIVDKLFYNAKMSLLDYLIVDPQSCNDDVRKDLTKLMAKGMTIQAAPRI 322
Db 61 PVLNFAIVDKLFYNAKMSLLDYLIVDPQSCNDDVRKDLTKLMAKGMTIQAAPRI 120

Qy 323 ROLLENLKLCAEVDNEMSLRTERHITFLDLCEENSLVYKVTGSDRGNAKYYDTLTF 382
Db 121 ROLLENLKLCAEVDNEMSLRTERHITFLDLCEENSLVYKVTGSDRGNAKYYDTLTF 180

Qy 383 KIYEENKKTFFPHPLPVKVGSGAKYAVPMHEHLEVEHEKQRYKNRIDLVMDKFLKRA 442
Db 181 KIYEENKKTFFPHPLPVKVGSGAKYAVPMHEHLEVEHEKQRYKNRIDLVMDKFLKRA 240

Qy 443 RKPHDYKENTLKMKELDPSSEELNVERFGLCSKLOMIECPGVKPEMLVNSVNOIK 502
Db 241 RKPHDYKENTLKMKELDPSSEELNVERFGLCSKLOMIECPGVKPEMLVNSVNOIK 300

Qy 503 MTPVIRGFQSKQLNVVPEKELCACAVFVYVNETAGNCPLEENDVVKFYTELIGGCKFRGIRI 562
Db 301 MTPVIRGFQSKQLNVVPEKELCACAVFVYVNETAGNCPLEENDVVKFYTELIGGCKFRGIRI 360

Qy 563 GANENRGAOSIMYDATKNEYAFYKNCITLNTGIGRFEIATAEAKNMFRLPDKEQVLMFI 622
Db 361 GANENRGAOSIMYDATKNEYAFYKNCITLNTGIGRFEIATAEAKNMFRLPDKEQVLMFI 420

Qy 623 IISKRQLNAGFYVKHYCDHTIGVANQHISETVTYKALASLRHEKSGKRIFYQIALKINAK 682
Db 421 IISKRQLNAGFYVKHYCDHTIGVANQHISETVTYKALASLRHEKSGKRIFYQIALKINAK 480

Qy 683 LGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVTHPTSYSGIDYSIAA VVASINPG 742
Db 481 LGGINQELDSEIAEISPEEKERRKTMPLTWYGVIDVTHPTSYSGIDYSIAA VVASINPG 540

Qy 743 GTTYRNMIVTQECRPOGERAVAGRETDILEAKFVKLLREFAENNDRAPAHIVVYRDG 802
Db 541 GTTYRNMIVTQECRPOGERAVAGRETDILEAKFVKLLREFAENNDRAPAHIVVYRDG 600

Qy 803 VSDSEMLRVSHDELRSLSKSVKQFMSERDGEDPEPKYTFVIQKRNTRLRLREWEKDKPV 862
Db 601 VSDSEMLRVSHDELRSLSKSVKQFMSERDGEDPEPKYTFVIQKRNTRLRLREWEKDKPV 660

Qy 863 VNKDLTPAETDVAVAQWEEEDMKESKETGIVNPSGGTTVDKLIYSKYKFDFFLASHHG 922
Db 661 VNKDLTPAETDVAVAQWEEEDMKESKETGIVNPSGGTTVDKLIYSKYKFDFFLASHHG 720

Qy 923 VLGTSRGCHYTYMTDDKMGSDQEVYKMTYGLAFISARCRKDISLPVPHVYAHLSCEKAKE 982
Db 721 VLGTSRGCHYTYMTDDKMGSDQEVYKMTYGLAFISARCRKDISLPVPHVYAHLSCEKAKE 780

Qy 983 LYRTYKEHYIGDYAQPRTREMEHFLQTNVVKYPGMSFA 1020
Db 781 LYRTYKEHYIGDYAQPRTREMEHFLQTNVVKYPGMSFA 818

RESULT 7
US-09-533-029-106
; Sequence 106 Application US/09533029
; Publication No. US20030046723A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omalra
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1149
US-09-533-029-106

Query Match 10.8%; Score 575.5; DB 3; Length 969;
Best Local Similarity 23.2%; Pred. No. 2.5e-34;
Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;

Qy 38 KKVLLLVNWKPSKSIYDREYIEYEVYKMTYKVLNRKPKGPPKTEIPIDRAKLFQHL 97
Db 158 KKVVRAN--HFLVQVADRDLHYDVSINPEVSK----- 190

Qy 98 RHEKKQTDFTLEDYVDEKDTVYSVCLNTVTGKMLVSEKVKKDS--KDEKDLKK 154
Db 191 -----TVNRNVKLLVN---YKDSHLGKSPAYD-GRK 220

Qy 155 ILVTM--ILTYRKXPHLNFSPENKDEANRSYK- LKNV-----MTQKRYA 200
Db 221 SLYTAGPLPDSKEFVYNLAEKADGSGGKDRPFKAVKNTVSTDLVQLQQLDRKQREA 280

Qy 201 PFVNEIKVQAKNFYDNNILRVDSFHDNRFQSL-----EVAPRIEAWFGI 251
Db 281 PY--DTIQVLDVVRDKPSNDYVSVGRSP-----PHTSLGKADRGELGDIYWRGY 333

Qy 252 YIGIKELFDGEPEVLNFAIVDKLFYNAKMSLLDYLIVDPQSCNDDVRKDLTKLMAK 311
Db 334 FQSLR-LTQWGLSLNIDVSARSFY-----EPVVTDIFISFLNIRD-L-N 375

Qy 312 MTRQAAARPRIROLLBNLKLCAEVDNEMSLRTERHITFLDLCEENSLVYKVTGSD-- 369
Db 376 RPLRSDRLKVKVKVRLTKVLLH-WNGTKSA-----KISGISLP 415

Qy 370 ----RGNNAKYYDTLTFKJYEENKFP-IEFPHPPLVVKVSGAKYAVPMHEHLEVEHEKQ 424

```
Db 416 IRELREFTLBDKSEKTVVQYFAEKYNYRVKYQALPAIQGSDTRPVVLPMLCQIDB-GQR 474
Qy 425 YKNRIDLVMQDFKPKRATKPKHDYKENTLMLKELDFSSSEELNVERFGLCSKQLMIECP 484
Db 475 YTKRLNEKQVTAALLKATCORPPD-RENSINKLVKNYNDLLS--KEFGMSVTTQLASTE 531
Qy 485 GKVLKEPMLVNSNEQIKWTPVIRGQEKQLNVVPEKELCCAVFVNVNAGNCPLEENDV 544
Db 532 ARVLPPMLKYHDSGKERNVPLG---QWNNIDKK----- 564
Qy 545 VKFYTEBLIGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCNTLNTGIRFEIAATEA 604
Db 565 -----MVNGAK---VTSWTCFQKQPAIPFISCPPEH-----IEEALLD- 600
Qy 605 KMFPERLPDKEQKVLMMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALASLRH 664
Db 601 --IHKEAPGLQ---LLVILPDVTGSGYKIKRICETELGIVSQCCQPRQVNL----- 648
Qy 665 EKGSKRIFYQIAIKNAKLGGINQELDWSEIAEISPEEKERRKTMPL-----TMYVGIDV 719
Db 649 ---NKQYMENVALKINVTGGRTVLN-----DAIRRNIPLIIDRPTIINGADV 694
Qy 720 THPTSYSGIDYSIAAVVASIN-PGGTIYRNMIWTO---EE-----CRPGERAVAHGR 767
Db 695 THPQGEDSSPSIAAVVASMDWPEINKYGLVSAQAHRHEIIQDLYKLVDPQGLVH-- 752
Qy 768 ERTDILEAKFVKLLREFAEANNDRAPAHIVVYRDGVSDSEMLRVSHDELSLKSEYKQFM 827
Db 601 --SGLIREHFIAPRA---TGQIPQRIIFYRDGVSEGFQSVLLHEMTAIRKACNSLQ 805
Qy 828 SERDGEDPBPKYTFVIOQRHNTRLRRMEKDKPVVKNKDLTPAETDVAVAQWBEEDMK 887
Db 806 -----ENYVPRVTFVIOQRHTRLRFPBOHG-----NRDMT----- 836
Qy 888 ESKETGINVPSSGTTVDKLIVSKYKDFDFLASHHGVLGTSRPGHYTYVMYDDKGMQDEVY 947
Db 837 --DKSGNIQP--GTVDVTKICHNEFDFFYLNHAGIQGTSRPAHYHVLDDENGFTADQLQ 892
Qy 948 KMTYGLAFLASRCRKPISLPVPVHYAHLSCAKELRYRKYKHYIGDYAQPRTR 1001
Db 893 MLTNNLCYTYARCTKSIVPPPAYAHAAFPRA---RYYMESEMSDGGSSRSR 942
```

RESULT 8

US-10-286-264-102

; Sequence 102, Application US/10286264

; Publication No. US20030093837A1

GENERAL INFORMATION:

; APPLICANT: Keddie, James

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Zhang, James

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Pineda, Omaisra

; APPLICANT: Heard, Jacqueline

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Adam, Luc

; APPLICANT: Broun, Pierre

; APPLICANT: Reuber, Lynne

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Samaha, Raymond

; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION

; FILE REFERENCE: MBI-008

; CURRENT APPLICATION NUMBER: US/10/286,264

; CURRENT FILING DATE: 2002-11-01

; PRIOR APPLICATION NUMBER: 60/125,814

; PRIOR FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 165

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 102

; LENGTH: 969

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

```
; FEATURE:
; OTHER INFORMATION: G1149
US-10-286-264-102
```

```
Query Match 10.8%; Score 575.5; DB 4; Length 969;
Best Local Similarity 23.2%; Pred. No. 2.5e-34;
Matches 235; Conservative 161; Mismatch 339; Indels 279; Gaps 40;
```

```
Qy 38 KVVLLVNWFKSSKIYDREYVEYKMTKEVLRKPKGPKFPKTEIPIDRAKLFWHL 97
Db 158 KKVVMYRAN--HFLVQVADRDLHYDVSNPEISK----- 190
Qy 98 RHEKQTDFILEDDYVDEKDTVYVCRNLNTVTSKMLVSEKVVKKDSE--KKDEKLEKK 154
Db 191 -----TVNRNVKLLVKN---YKDSHLGGKSPAYD-GRK 220
Qy 155 ILYTM--ILTYRKPKPLNFSRENPEKDEANRSYKF-LQNV-----MTQKRVYA 200
Db 221 SLYTAGPLPFDSEKFEFVNLAEKRADGSSGKDRPFKVVAVKNVTSTDLYQLQQFLDRKREA 280
Qy 201 PFVNEEIKVQFAKNFYVDNNSILRVPESPHDNPRQEQL-----EVAPEIEAWFGI 251
Db 281 PY--DTIQVLDVVLDRKPSNDYVSVGRSF-----PHTSLGKDARDGRGLGDIYWRGY 333
Qy 252 YIGIKELPDGEPVLPFAIVDKLPYNAPKMSLLDYLILLIVDPOSNDNDVRKDLKTKLMAGK 311
Db 334 FQSLR-LTQMGLSLNIDVSARSFY-----EPIVVTDFISKFLNIRDL--N 375
Qy 312 MTIROAARPRIROLLLENKLKCAEVDNMSRLTERHLTFLDLCSENSLVYKVTGKSD-- 369
Db 376 RPLRSDRLKVKVKVLRTLKVKLLH-WNGTKSA-----KISGSSLP 415
Qy 370 ----RGRNAKKYDTTLFKIYEENKCF-IEFPHLPVVKVSGAKAYAVPMHELVHEKPKOR 424
Db 416 IRELFTLEDSEKTVVQYFAEKYNYRVKYQALPAIQGSDTRPVVLPMLCQIDB-GQR 474
Qy 425 YKNRIDLVMQDFKPKRATKPKHDYKENTLMLKELDFSSSEELNVERFGLCSKQLMIECP 484
Db 475 YTKRLNEKQVTAALLKATCORPPD-RENSINKLVKNYNDLLS--KEFGMSVTTQLASTE 531
Qy 485 GKVLKEPMLVNSNEQIKWTPVIRGQEKQLNVVPEKELCCAVFVNVNAGNCPLEENDV 544
Db 532 ARVLPPMLKYHDSGKERNVPLG---QWNNIDKK----- 564
Qy 545 VKFYTEBLIGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCNTLNTGIRFEIAATEA 604
Db 565 -----MVNGAK---VTSWTCFQKQPAIPFISCPPEH-----IEEALLD- 600
Qy 605 KMFPERLPDKEQKVLMMFIIISKRQLNAYGFVKHYCDHTIGVANQHITSETVTKALASLRH 664
Db 601 --IHKEAPGLQ---LLVILPDVTGSGYKIKRICETELGIVSQCCQPRQVNL----- 648
Qy 665 EKGSKRIFYQIAIKNAKLGGINQELDWSEIAEISPEEKERRKTMPL-----TMYVGIDV 719
Db 649 ---NKQYMENVALKINVTGGRTVLN-----DAIRRNIPLIIDRPTIINGADV 694
Qy 720 THPTSYSGIDYSIAAVVASIN-PGGTIYRNMIWTO---EE-----CRPGERAVAHGR 767
Db 695 THPQGEDSSPSIAAVVASMDWPEINKYGLVSAQAHRHEIIQDLYKLVDPQGLVH-- 752
Qy 768 ERTDILEAKFVKLLREFAEANNDRAPAHIVVYRDGVSDSEMLRVSHDELSLKSEYKQFM 827
Db 753 --SGLIREHFIAPRA---TGQIPQRIIFYRDGVSEGFQSVLLHEMTAIRKACNSLQ 805
Qy 828 SERDGEDPBPKYTFVIOQRHNTRLRRMEKDKPVVKNKDLTPAETDVAVAQWBEEDMK 887
Db 806 -----ENYVPRVTFVIOQRHTRLRFPBOHG-----NRDMT----- 836
Qy 888 ESKETGINVPSSGTTVDKLIVSKYKDFDFLASHHGVLGTSRPGHYTYVMYDDKGMQDEVY 947
Db 837 --DKSGNIQP--GTVDVTKICHNEFDFFYLNHAGIQGTSRPAHYHVLDDENGFTADQLQ 892
Qy 948 KMTYGLAFLASRCRKPISLPVPVHYAHLSCAKELRYRKYKHYIGDYAQPRTR 1001
```


Db 893 MLTNLCYTYARCTKSIVPPAYAHAAAFRA-----RYTWESEMSDGGSSRSR 942

RESULT 9

US-10-374-780A-2084

; Sequence 2084, Application US/10374780A

; Publication No. US20040019927A1

GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James

; APPLICANT: Broun, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10/374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR APPLICATION NUMBER: 09/837,944

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 10/225,066

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,067

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,068

; PRIOR FILING DATE: 2002-08-09

; NUMBER OF SEQ ID NOS: 2906

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2084

; LENGTH: 969

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; OTHER INFORMATION: G1149 Paralogous to G1146

US-10-374-780A-2084

Query Match 10.8%; Score 575.5; DB 4; Length 969;

Best Local Similarity 23.2%; Pred. No. 2.5e-34;

Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;

QY 38 KKVLLVWPKFSKSLYDREYVEYVMTKEVLRKPKFPKTEIPIDRAKLFQHL 97

Db 158 KKVWVRAN--HFLVQVADRDLHYDVSINPEVISK----- 190

QY 98 RHEKKQTDFTLEDYVFEKDTVYSVCKLNTVTSMKLVSEKVKVKKDSE---KKDEKLEKK 154

Db 191 -----TVNRNVKLLVKN---YKSHLGKSPAYD-GRK 220

QY 155 ILVTM--ILTYRKKFHLNFSRENPKDEANRSYKF-LKNV-----MTQKRVYA 200

Db 221 SLYTAGPLPDSKEFVNVLAEKRAKDGSGKDRPFKVAKNVTSTDLVQLQQLDRKQEA 280

QY 201 PFVNEIKVQPAKNFYVDNNSILRVPESHDPNRFQSL-----EVAPRIEAWFGI 251

Db 281 PY--DTIQVLVDVLRDKPSNDYVSVGRSP-----FHTSLGKARDRGELGDGIEYWRGY 333

QY 252 YIGIKELFDGEPVLNPAIDKLYNAPKMSLLDYLILLVDPSCNDVDVRLDKTKLMAGK 311

Db 334 FQSILR-LTQMGLSLNIDVSARSFY-----EFIVVTDIFISKFLNTRDL--N 375

QY 312 MTIRQAARPRIQLLENLKLKCAEVWDNEMSLRTERHLTLFDLCEENSLVYKVTGKSD-- 369

Db 376 RPLRDSRLKVKVLRTLKVKLLH-WNGTKSA-----KISGISLPL 415

QY 370 ----RGRNAKYDTTLFKIYEENKCP-IEPPLPLVVKVSKAKAYAVPMHELVHEKQOR 424

Db 416 IRELFTLEDKSEKTVQVFAEKYNRVKYQALPAIQTGSDTFRVPLFMELCQIDE-QOR 474

QY 425 YKNRIDLVMDQKFLKRAIRKPHDYKENTLKMELDFSEBELNVEFRFGLCKLQMBICP 484

Db 475 YTKELNEKQVTAALKATCQRPDP-RENSIKLVVKNYNDLDS--KEFGHSTVTLASIE 531

QY 485 GKVLKEPMLVSNVEIQIKMTFVIRGFQEKQLNVVPEKELCCAVFVNNETAGNPLENDV 544

Db 532 ARVLPPEMLKYHDSGKEKMNPRLG---QWNMDKK----- 564

QY 545 VKTYTELIGCKFRGIRIGANENRGAQSIMYDATKNEYAFYKNCNTLNTGIRFPRIATEA 604

Db 565 ----MVNGAK--VTSWTCBFKQPAIPFISCPPEH-----IEALLD- 600

QY 605 KMPFERLPDKEQKVLAFIISKQLNAYGFVKHYCDHTIGVANOHITSETVTKALASLRH 664

Db 601 --IHKAPGLQ----LLIVLPDVTGSYGKIKKICETELGIVSQCCQPRQVNL- 648

QY 665 EKGSKRIFYQIALKINAKLGGINOELDWSIABISPEEKERRKTMPL-----TMYVGIDV 719

Db 649 ---NKQVMENVALKINVKTGGRNTVLN-----DAIRRNIPLTIDRPTIIMGADV 694

QY 720 THPTSYSGIDYSTAAVVASIN--PGCTIYRNMIYTO---EE-----CRPGERAVAHGR 767

Db 695 THPQGEDSSPSIAAVVASMDWPEINKYRGLVSAQAHRSEIIQDLYKLVQDPQRLVH-- 752

QY 768 ERTDILEAKFKVLLREFAEENNDRAPAHIVVYRDGVSDSEMLRVSHDELRLSKSEVKQFM 827

Db 753 --SGLIREHPIAFRA-----TCQIPQRIIFYRDGVSEGFQSVLLHEMTAIRKACNSLQ 805

QY 828 SERDGEDPPEKYTFIVIQKRNHTRLLRRMEKDKPVVNVKDLTPAETDVAAVAKQWEEDMK 887

Db 806 ----ENYVPRVTFVIVQKRHTLFPQHG-----NRDMT----- 836

QY 888 ESKETGVNPSGTTVDKLVSKYKDFDFFLASHHVLGTSRPGHYTMYDDKMGMSQDEVY 947

Db 837 --DKSGNIQP--GTVVDTKICHENEFDFYLNSHAGIQGTSRPAHYHVLLENGFTADQLQ 892

QY 948 KMTYGLAFLSARCKRPISLPVPVHYAHLSCAKELYRTYKEHYIGDYAQPRT 1001

Db 893 MLTNLCYTYARCTKSIVPPAYAHAAAFRA-----RYTWESEMSDGGSSRSR 942

RESULT 10

US-10-412-699B-752

; Sequence 752, Application US/10412699B

; Publication No. US20040045049A1

GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Zhang, James

; APPLICANT: Fromm, Michael E.

; APPLICANT: Heard, Jacqueline E.

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Adam, Luc J.

; APPLICANT: Broun, Pierre E.

; APPLICANT: Pineda, Omaira

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James S.

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Jiang, Cai-Zhong

```
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 752
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1149
; US-10-412-699B-752

Query Match          10.8%; Score 575.5; DB 4; Length 969;
Best Local Similarity 23.2%; Pred. No. 2.5e-34;
Matches 235; Conservative 161; Mismatches 339; Indels 279; Gaps 40;

QY 38 KVVLLVNNKFSKKIYDREYEVYKWKVNLNKKPKPKPKTEIPIDRAKLFWQHL 97
DB 158 KKVVRAN--HFLVQVADRDLHYDVSINPEVISK----- 190
QY 98 RHEKQTDFILEDYVDEKDTVYVSVCLNTVTSKMLVSEKVVKKDSE---KKDEKDKLEKK 154
DB 191 -----TVNRNVNKKLLVKN---YKSHLGKSPAYD-GRK 220
QY 155 ILYTM--ILTYRKKFHLNFSRENPERDEANRSYKF-LKNV-----MTQKRVYA 200
DB 221 SLYTAGPLPDSKEFVNNLAERADGSSKDRPFKVAKNVSTDLVQLQQFLDRKQREA 280
QY 201 PFVNEIKVQFNAKFNVDNNSILRVPESHDNRPEQSL-----EVAPRIEAWFGI 251
DB 281 PY--DTIQVLDVLRDKPSNDYVSVGRSF-----FHTSLGKDARDGELGDGIEYWRGY 333
QY 252 YIGIKELFDGEPVNLNPAIVDKLFPYNAPKMSLLDYLILLIVDPQSCNDVDVRKDKTKLMACK 311
DB 334 FOSLR-LTQWGLSLNIDVSARSFY-----EPIVVDTFISKFLNIRDL--N 375
QY 312 MTRQARPRIRQLLENLKLKAEVWDMNMSRLTERHLTFLDLICEENSLVYKVTGKSD-- 369
DB 376 RPLRDSRLKVKVLRATLKVLLH-WNGTKSA-----KISGISSLP 415
QY 370 ----RGRNAKVDITLTKIYENKCF-IEFPHLPLVKVSGAKYAVPWEHLEVHEKPOR 424
DB 416 IRELFTLEDKSEKTVVQYFAEKYINRVYQALPAITGSDTRPVVLPMLCQIDE-GOR 474
```

```
QY 425 YKVRIDLVMQDKFLKRAIRKPHDYKENTIKMLKELDFSEELNVERFGLCSKLOMIECP 484
DB 475 YTKLNEKQVNTALLKATCORPPD-RENSISKLVNKNYNDLUS--KEFGMSVTTQASIE 531
QY 485 GKVLKEPMLVNSVNEQIKMTVPVIRGFBQKQLNVNPEKELCCCAVFFVNVNETAGNPCLEENDV 544
DB 532 ARVLPPMLKYHDSCKEKWVNERLG---QNMWDKK----- 564
QY 545 VKFYTELGGCKFRGIRIGANENRGAQSIWYDATKNYAFYKNCNTLNTGIGRFEIAATEA 604
DB 565 -----MVNGAK---VTSWTCFKPQPAIPFISCPPEH-----IEEALLD- 600
QY 605 KMFPERLPDKEQKVLMMFIIISKRLNAYGFVGHYCDHTIGVANQHITSETVTVKALASLRH 664
DB 601 --IHKRAFGQ-----LLVILPDVTGSGYKIKRI CETELGIVSQCCQPRQVKNL----- 648
QY 665 EKGSKRIIFYQIALKINAKLGGINBELDWSEIAEISPEEKERRKTWPL-----TMYYGIDV 719
DB 649 ---NKQYMNVALKINVKTGGRNTVLN-----DAIRNTPLITDRPTIIMGADV 694
QY 720 THPTSYSGIDYISIAAVASIN--PGGTIYRNMIWQ---EE-----CRPGERAVAHGR 767
DB 695 THPQGEDSSPSIAAASVMSMDWPEINKYRGLVSAQAHRBEIIODLYKVQDPQRGVLH-- 752
QY 768 BRTDILEAKFVKLLREFAENNDRAPAHIVVYRDGVSDEMLRVSHDELRLSKSEVKQFM 827
DB 753 --SGLIREHFTAFRA-----TGQIPQRIIFRDGVSSEGQFQVOLLHEMTAIRKACNSLQ 805
QY 828 SERDGEDPEPKYFTFIVIOQRNHTLLRMEKDKPVVKNDLTPAETDVAVAVKQWBEEDMK 887
DB 806 -----ENVYPRVTFVIVQKRHHTRLPPEQHG-----NRDMT----- 836
QY 888 ESKETGINVPSSGTTVDKLIIVSKYKFDPLASHHGVLTGTSRPHYHYVMYDDKMSODEVY 947
DB 837 --DKSGNIQP--GTVDYTKICHNPBDFYLNSHAGIQGTSRPAHYHVLVDENGFTADQLQ 892
QY 948 KMTYGLAFLASRCRKPISLPVVPVVAHLSECAKELYRKYKEHYIGDYAQPRTR 1001
DB 893 MLTNNLCYTYARCTKSIVPPAYVAHLAAFPRA----RYMESEMSDGGSSRSR 942

RESULT 11
US-10-467-397-7
; Sequence 7, Application US/10467397
; Publication No. US20040137448A1
; GENERAL INFORMATION:
; APPLICANT: THORNTON, Michael; HAFALIA, April J. A.;
; APPLICANT: LU, Dyung Aina M.; ARVIZU, Chandra S.;
; APPLICANT: SWARNAKAR, Anita; LU, Yan;
; APPLICANT: WARREN, Bridget A.; BAUGHN, Maria R.;
; APPLICANT: TANG, Y. Tom; LEE, Ernestine A.;
; APPLICANT: YAO, Monique G.; RAMKUMAR, Jayalaxmi;
; APPLICANT: KHAN, Farrah A.; GANDHI, Ameena R.;
; APPLICANT: DING, Li; YUE, Henry;
; APPLICANT: GIETZEN, Kimberly J.; CHAWLA, Narinder K.;
; APPLICANT: THANGAVELU, Kavitha; ELLIOTT, Vicki S.
; APPLICANT: MARQUIS, Joseph P.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PI-0368 USN
; CURRENT APPLICATION NUMBER: US/10/467,397
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US02/03844
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/268,118
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/270,963
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,858
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/271,194
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,071
```

```

; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/283,496
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/344,650
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1725129C01
US-10-467-397-7

```

Query Match	10.6%;	Score 566;	DB 4;	Length 860;
Best Local Similarity	23.1%;	Pred. No. 1.1e-33;		
Matches	234;	Conservative 160;	Mismatches 370;	Indels 250; Gaps 39;
QY	19	PENKMLARP	CGKDGKFEYK	KVLLVNWPFESSKIYDREYVEYVMTKEVLRNPKGPP 78
DB	18	PIGYAFKPPRPDP	FGSGRTIKLQANFFEMDIPKIDIVYELDIKPEK	-----C 67
QY	79	PKTEIPI	DPRAKLFW	OHLEHHEKQTDFILEDY--VPDEKDTVYVSCVRLNLTVTSKMLVSE 136
DB	68	PRVNR	EIVEHM-----VQHEFTQ	---IFGDRKPFVDGRKNLYTAMPLPIGRDKV--E 115
QY	137	KUVKQSEKDEK	DEKLEKILYTMILYTRKXPHLNF	SRENPEKDEANRSYKFLKNVMTQ- 195
DB	116	LEVTLP	CEGKOR--IFKVSIKW	SCVSLQALHDAISGRLPSPPE---TIQALDDVVRHIL 170
QY	196	-KRVYAPFVNEEIK	VQAKNFVYDNNISILRVPSFHD	PNRPEQSLEVAPRIEAWFGVIYG 254
DB	171	PSNRYPV	-----GRSFFTASEGCSNP	-----LGGREWVFGFHQS 206
QY	255	IKELF	CGEPVLNPAIVDKLPYNAPKMSLLD	YLLIVDPQSCNDDVRKDLKTKMLAGKWTI 314
DB	207	VRPSL-WQWMLN	IVDSATAFYKA-QPVIEFVCEVLD	FKSIEEQ-----QKPL 251
QY	315	ROAARPR	IROLLENLXK--CAEVDN	----EWSRLTERHLPTDLCEENSLVYKVTGK 367
DB	252	TUSQRVKFTKEI	IKGLKVEITHCCQMKRYKRV	CNVTTRPASHQTF-PLQSSGQTVECT-- 308
QY	368	SDRGNAKYD	YDTTLFKIYEENKEFIEFP	PHLPVLRKVSKAKEYAVPMHEHLEVHEKQPXYKN 427
DB	309	-----VAQY	-----FK--DRHKLVLRYPHLP	CLVQVGEQKHVTLFLEVCNIVAGORCIKK 356
QY	428	RIDLVMQDF	ELKRAT-RKPHDYKENTL	KMLKELDFSEELNLFVERFGLCSKLQWIECPGK 486
DB	357	LTD--NOT	STMIRATARSAPDROGEET	SKMRSASFNTDP--YVREFGIMVKDEMTDVTGR 412
QY	487	VLKEPMLVNSV	BEIQIKMTVP--IRGFOEKQLNV	-----VPEKELCCAVFVUNE 532
DB	413	VLQPPSILY	GGRNKATATPVQGVWDMENK	QFHTGIELKWALACFAPQO-CTEVHLKSP 471
QY	533	T-----	AGNPCLENDVVKFTYELIG	CGCKFGIRIGANENRGAOSI--MYD 576
DB	472	TEQLRKIS	DAGMPIQGP	CF-----CKYA-----QGADSVEEMFR 507
QY	577	ATNEVAFYKNC	TLNTIGRFEIATAEAKNMPER	LPDKEQVLMFIILSKROLNAYGFVK 636
DB	508	HLKNTYA	-----GLQVVLIPGKTPV	YAEVK 534
QY	637	HYCDHTIGV	ANQHITSETVTKALASLR	HEKGSKRIFVQIALKINAKLGGINQBELWSEIA 696
DB	535	RVGDTVL	GMATQCVQMKNV-----QRTTPTQLSNL	CKINVKLGGVNNIL----- 579
QY	697	EISPEKERR	KTKMPLTWYIGIDVTHPT	SYSGIDYSTAAVVASINPGGTIYRNIMVTEEC 756
DB	580	--LPQRP	PPVQFPV-IFGLADVTHP	PAGGKKPSAAVVVGSMDAHPNRYCATRVQO-- 634
QY	757	RCGERAVAH	GRERTDILEAKFVKL	REFAENNDRAPAHIVYVRDGVSDSEMLRVSHDEL 816

635	Db	-----HQEIIQDLAAWVRELLIQFYKSTRFK-PTRIIFYRDGVSEGOQQVLHHEL	685
817	Qy	RSLKSEVKQFMSERDGEDPEPKYTFIVIQKHNTLLRRMEKDKPVNKDLTPAETDVAV	876
686	Db	LAIREACIKL-----EKDYQPGITFIVVQKHHTLRF-CTDKNERVGKSGNIPA-----	733
877	Qy	AAVQWEEDMKESKETGIWNPSSGTTVDKLVSKYKDPFPFLASHHGVLGTSRPGHYTMY	936
734	Db	-----GTTVDTKITHPTFPDFYLCSHAGIQCTSRPSHYVLW	770
937	Qy	DDKGMODEVVKMTYGLAFISARCKEPTSLPVPVHYAHLSEKAKELRTYIEKH	990
771	Db	DDNRFSDELQILQYLCHTYVTRTSVTSIPAPAYIAHLVAFRAE-VHLVDKEKH	823

RESULT 12

```

RESULTS 12
US-11-097-143-27930
; Sequence 27930, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27930
; LENGTH: 950
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-27930

```

Query Match	10.5%;	Score	560.5;	DB	6;	Length	950;
Best Local Similarity	22.9%;	Pred. No.	3.4e-33;				
Matches	225;	Conservative	170;	Mismatches	386;	Indels	201;
Gaps	35;						
QY	38	KKVLLLVNMFKSSKIYDREY-VEYEVKMTKEVLNRKPGKPPPKTEIPIPDRAKLPWQH	96				
DB	112	RIVLEAHHFQVT---MPRGYVHHYDINTPQCKPKVNRRI---LETWVHYSKIFG--	163				
QY	97	LAHEKKQTDFILEDYVDFDEKOTVYSVCRLLNTVTSKMLVSEKVKVKDSEKKDKLEKKIL	156				
DB	164	-----VLKP-VFDGRNNLYTRDPLPIGNERL---ELEVTLPGEQKDR-----	201				
QY	157	YTMILTYRKPHLNTSRENPEKDEBANRSYKFLKNVMTOKVRYAPVPNRSEIKVQFAKNFV	216				
DB	202	---IFRVTIKQAQVSLFMLEALEGR-----TRQIPYDAITALDVVNRHLPMSMT	248				
QY	217	YD--NNSILRVPESHDPNRFQSLUEVAPRIEAWFGIYTGIKELFQGEVPLNFAIVDKLF	274				
DB	249	YTFVGRSFPSSPEGYYHP-----LGGGREGVWFGFHQSVRP-SQWQMLNIDVSGATAP	299				

```

QY 275 YNAPKMSLLDYLLLVDPQSCNDDVRKDLTKLTKLMAGKMTIROAARPRIRQLLENLKLKCA 334
Db 300 YKA--QPVIDFMCVLDIIRINEQ-RKPL-----TDSQRVKFTKEIKGLKIEIT 345
QY 335 EVDNEMSLRTERHLTFLDLCENSLVYKVTGKSDRG-----RNAKYDYDTTLFKI 384
Db 346 H-----CGQMRKRYVCNVRTPAQMSFPPLQLENGQIVECTVAKY 386
QY 385 Y-BENKKEIEPPLPLVVKVSKAKAYAVPMHELVHEVKPORYKNRIDLVMDQKFLKRAFR 443
Db 387 FLDKRYMKLRYPHLPCLQVQEHKHTYLPLEVNI--VAGQRCIKKLTDMQJSTMIKATAR 445
QY 444 KPHDYKENTLKMELDFSEELNFERFGLCSKLQMIKPCGKVLKEPMLVNSVNEIQIM 503
Db 446 SAPDREREINLVKRAADFND--SYQBFGLTISMSMEVGRVLPPLPKLOYGGR----- 498
QY 504 TPVIRGFBQKQNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGGCKFRGIRIG 563
Db 499 --VSTGLTGQQL-FPPQNKVSLA-----SPNQGVMDMRGKQFFTG-----EIRIW 541
QY 564 ANENRGAQ-SIMYDATKNEYAFYKNCNTLTG---IG-----RPEATAAKMFERLPDK 614
Db 542 AIACFAPQRTVREDALRNFTQQLKISNDAGMPIIGQPCFCYATGPDQVPEMFYRLKIT 601
QY 615 EQKVLMIISIKROLNAYGFVKHYCDHTIGVANQHITSETVTKALASLRHEKSGKRIFYQ 674
Db 602 FPGQLVVLVLPCKTPVYAEVVRVGDVTLGMATQCQVAKNVKT-----SPQTLN 652
QY 675 IALKINAKLGGINQELDMSEIAEISPEEKERRKTWPLTWYGVDTVHTPTSYSGIDYSIAA 734
Db 653 LCLKINVKLGGINSIL---VPSIRPKVFNEP-----VIFLGADVTHPPAGDNKKPSIAA 703
QY 735 VVASINPGGTIYRNMIIVTQEECPGERAVAHGERTDILEAKFVKLLRFAENNDRAPA 794
Db 704 VVGSDAHPSPRYAATVRVQQ-----HRQEIQLSSVMVRELLIMFYKSTGGYKPH 753
QY 795 HTVYRDGVSDEMLRVSHDELRSLSKSEVKQPMSERDGEDPE--PKYTFIVIQKHNTRL 852
Db 754 RLILYRDGVSEGFPHVLQHELTATREACIKL-----EPEYRGITFIVVQKHHTL 806
QY 853 LRMEKDKPVVKNLTPAETDVAVAVKQWEEDMKESKETGVNPSSTGTTVDKLVSKYK 912
Db 807 FCAEKE-----QSGKSG--NIPAGTTVDVGITHPTE 836
QY 913 FDFFLASHGVLTSPRGHTYMYDDKMGSDQSEVYKMTYGLAFLSARCKPISLVPVHY 972
Db 837 FDFYLCSHQGIQGTSPSHYHLVMDNDHDFDSELOCLTYQLCHTYVVRCTRSVSIPAPAY 896
QY 973 AHLSCSEKAKELYRTYKEHYIGD 994
Db 897 AHLVAFRAR-YHLVEKEHDSGE 917

```

RESULT 13

```

US-11-097-143-14511
; Sequence 14511, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769

```

```

; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14511
; LENGTH: 984
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-14511

Query Match      10.5%; Score 560.5; DB 6; Length 984;
Best Local Similarity 22.9%; Pred. No. 3-6e-33;
Matches 225; Conservative 170; Mismatches 386; Indels 201; Gaps 35;

QY 38 KKVLLLVNMPKSPSSKIYDREY-YEYEVKMTKEVLNARKPGKPPKTEIPIPRAKLFWOH 96
Db 146 RPVLNANHFQVT--MPRGYVHHYDINIQDPKCPKVNREI---IETVHVHAYSKIFG-- 197
QY 97 LRHEKKQDTPFILEDYVDEKDTVYSVCLNTVTYSKMLVSEKVVKKDSEKKDEKLEKKIL 156
Db 198 -----VLKP-VFQGRNNLYTRDPLPIGNERL-----ELEVTLPGEGKDR----- 235
QY 157 YTMILTYRKFKHLNFSRENPEKDEBANRSYKFLKGNVMTOKRYAPFVNEEIKVQFAKNFV 216
Db 236 ---IFRVITKWAQVSLFNLEALEGR-----IRQIPYDAILALDVMVHRHLSMT 282
QY 217 YD--NNSILRVPESHDNRPEQSLEAVPRIAWEFGIYIGIKELFDGEPVLPAIVDKLF 274
Db 283 YTPVGRSPFSPSEGYHP-----LGGGREGVWFQHQSVRP-SQWKMLNIDVSATAF 333
QY 275 YNAPKMSLLDYLLLVDPQSCNDDVRKDLTKLTKLMAGKMTIROAARPRIRQLLENLKLKCA 334
Db 334 YKA--QPVIDFMCVLDIIRINEQ-RKPL-----TDSQRVKFTKEIKGLKIEIT 379
QY 335 EVDNEMSLRTERHLTFLDLCENSLVYKVTGKSDRG-----RNAKYDYDTTLFKI 384
Db 380 H-----CGQMRKRYVCNVRTPAQMSFPPLQLENGQIVECTVAKY 420
QY 385 Y-BENKKEIEPPLPLVVKVSKAKAYAVPMHELVHEVKPORYKNRIDLVMDQKFLKRAFR 443
Db 421 FLDKRYMKLRYPHLPCLQVQEHKHTYLPLEVNI--VAGQRCIKKLTDMQJSTMIKATAR 479
QY 444 KPHDYKENTLKMELDFSEELNFERFGLCSKLQMIKPCGKVLKEPMLVNSVNEIQIM 503
Db 480 SAPDREREINLVKRAADFND--SYQBFGLTISMSMEVGRVLPPLPKLOYGGR----- 532
QY 504 TPVIRGFBQKQNVVPEKELCCAVFVNNETAGNCPLEENDVVKFYTELIGGCKFRGIRIG 563
Db 533 --VSTGLTGQQL-FPPQNKVSLA-----SPNQGVMDMRGKQFFTG-----EIRIW 575
QY 564 ANENRGAQ-SIMYDATKNEYAFYKNCNTLTG---IG-----RPEATAAKMFERLPDK 614
Db 576 AIACFAPQRTVREDALRNFTQQLKISNDAGMPIIGQPCFCYATGPDQVPEMFYRLKIT 635
QY 615 EQKVLMIISIKROLNAYGFVKHYCDHTIGVANQHITSETVTKALASLRHEKSGKRIFYQ 674
Db 636 FPGQLVVLVLPCKTPVYAEVVRVGDVTLGMATQCQVAKNVKT-----SPQTLN 686
QY 675 IALKINAKLGGINQELDMSEIAEISPEEKERRKTWPLTWYGVDTVHTPTSYSGIDYSIAA 734
Db 687 LCLKINVKLGGINSIL---VPSIRPKVFNEP-----VIFLGADVTHPPAGDNKKPSIAA 737
QY 735 VVASINPGGTIYRNMIIVTQEECPGERAVAHGERTDILEAKFVKLLRFAENNDRAPA 794
Db 738 VVGSDAHPSPRYAATVRVQQ-----HRQEIQLSSVMVRELLIMFYKSTGGYKPH 787

```

```

QY 795 HIVYRDGVSEMLRVSHDELRSLSKSEVKQPMSERDGEDPE--PKYTFIVIOKHNTRL 855
Db 788 RIILYRDGVSEGGQFPHVLQHELTAREACIKL-----EPEYRPGITFIVVQKGRHTRL 840
QY 853 LRRMEKDPVNVKDLTPAETDVAVAAVQWESDMKESKETGIVNPSGGTTVDKLVSKYK 912
Db 841 FCAEKK-----QSKSG--NIPAGTTVDVGTHTPTE 870
QY 913 PDDFLASHHGVLTGSRPGHYTVMYDDKGMQSDVEYKMTYGLAFLSARCKPISLPPVPHY 972
Db 871 PDFVLCSHQGIQGTSRPSHYHVLWDDNHFDSDQLCLTVQLCHTVVRCRSVSIPAPAY 930
QY 973 AHLSCERAKELYRYTYKEHYGD 994
Db 931 AHLVAFRAR-YHLVEKEHDSGE 951

RESULT 14
US-11-097-143-27927
; Sequence 27927, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27927
; LENGTH: 984
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-27927

```

[illegible]

Qy	275	YNAPENSLDYLLIVDPQSCNDVDRKDLKTKLMAGKWTIRQAAAPRTRQLLENLKLKCA	3334
Db	334	YKA--QPWIDPFCVLDIRNEQ-RKPL-----TDSQRVKFTKEIGLKIET	379
Qy	335	EVWDNEMSRLTERHLTFLDLCBENSUVKVTGKSDRG-----ENAKKYDTTLPKI	384
Db	380	H-----CGWRKRVNCVTRRPAQMOSFPLOLENGQVTECTVAKY	420
Qy	385	Y-EENKKFIEPPHPLPLVKVSGAKAYAVPMHELVHEKPORYKNRIDLVMODKFLKRA	443
Db	421	FLDKYRMKRLRYPHLPCLQVQGEKHHTYLPLEVCNI-VAGQRCIKKLTDMQWSTMIKATAR	479
Qy	444	KPHDYKENTLKMELDEPSSSELNPFVERFGLCSKLQMIETCEGKVLKEPMLVNSVNEQIM	503
Db	480	SAPDREREINNLVKRADFNND--SYVQEFGLTISNSMMEVGRVLLPPKQLQYGR	532
Qy	504	TPVIRGFEQKQLNVVPEKELCCAVFVNNETAGNCLBENDVVKFVTELIGGCKPGRIG	563
Db	533	--VSTGLTGQOL-FPPQNKVSLA-----SPNGVMDMRGKQFPTGV-----EIRIW	575
Qy	564	ANENRGAQ-SIMYDQATKNEYAFYKNCNTLNTG---IG-----RFEIATAEAKNMFERLPDK	614
Db	576	AIACFAQRTVREDALRNFTQOLQKISNDAGMPILGQPCFKYATGPDQVEPMFYLKIT	635
Qy	615	EQVLMFIITISKQLNAYGVFKHYCDHTIGVANQHIITSETVTYKALASLRHEKSGKRI	674
Db	636	PPGLQLVVVVLPKGTPTVYAEVKRYGDTVLGMATOCVQAKNVNKT-----SPQTL	686
Qy	675	IALKINAKLGINQELQDSEIABEISPEEKERRKTMTMYVGIDVTHPTSYSGIDYSTAA	734
Db	687	LCLKINVKLGINSIL-----VPSIRPKVNEP-----VIFUGADVTHPPAGNKKPS	737
Qy	735	VVASINPGGTIYRNMIVTQBECPGERAVAHGREBTDILEAKFVKLRFEFANENDNRAPA	794
Db	738	VUGSMDAHPRYAATVRVQ-----HRQEIQLSELSMWRELLIMFYKSTGGYKPH	787
Qy	795	HIVYRQGVSDSEMLRVSHDELBSLSEVKQFMSERGEDBE--PKYTFVIQKRNHTRL	852
Db	788	RIILYRDGVSGEGOPPHYLQHELTAIREACINL-----EPEYRPGITFIVVQKRHH	840
Qy	853	LRRMEKDKPVVKNDLTPAETDVAAVVKQWEEDMKESKETGIVNPSGTTVDKLTIVSKV	912
Db	841	FCAEKK-----QSGKSG--NIPAGTTVDVGITHPTE	870
Qy	913	FDFFLASHHGVLGTSRPGHYTVMYDDDKGMSQDEYVKMYTIGLAFLSARCKRKPISLP	972
Db	871	FDYFLCSHQIGQTSRPSHYHVLWDDNHFDSDLOCLTYQLCHTYVTRTSVSPAPAY	930
Qy	973	AHLSCEKAKBLRYTKBHYTGD	994
Db	931	AHLVAFRAR-YHLVKEKHDGSGE	951

RESULT 15
US-10-645-746-9
Sequence 9, Application US/10645746
Publication No. US20040265839A1
GENERAL INFORMATION:
APPLICANT: Mello, Craig C.
APPLICANT: Tabara, Hiroaki
APPLICANT: Grishok, Alla
APPLICANT: Fire, Andrew
TITLE OF INVENTION: RNA INTERFERENCE
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: UMY-052DV1
CURRENT APPLICATION NUMBER: US/10/645,746
CURRENT PILING DATE: 2003-08-20
PRIOR APPLICATION NUMBER: US 09/689,992
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/193,218
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US 60/159,776

```

; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-645-746-9

Query Match      10.4%; Score 557; DB 5; Length 766;
Best Local Similarity 23.4%; Pred. No. 4.5e-33;
Matches 184; Conservative 139; Mismatches 312; Indels 150; Gaps 23;

QY 230 HDNPEQSLVAPRIEAWFGIYIGIKELFDGEPVNFPAIVDKLFYNAPKMSLLDYLLLI 289
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
89 HSAGQTHAESKLGGRGVWFGHQSVRP--SOWKMLNIDVSATAFYRS--MPVIEFIAEV 145
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 VDPQSCNDVDRKDKTKLMAGKMTIROAARPRIROLLENLKLKCAEVWDMNEMSLRTERHL 349
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 L-----ELPQALAEERRALSDAQRVKFTKEIRGLKIEITH----- 180
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 TFLDLCEENSLVYKVTGSDGRNKKY-----DTTLPK-IYEENKKPIBFPHP 398
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 -----CGOMRRKYRVCNVRTTRPAQTQTPLEQTGTIECTVAKYFYDKYRIQLKYPHLP 235
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
399 LVYKSGAKAYAVPMHEHLEVEHVKPQRYKNRIDLVMDKFLKRAKTRKPHDYKENTLMLKE 458
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236 CLOVGOEQKHLYLPPEVCNI-VPGQRCKIKKLTVDVSTMTIKATARSAPEREREISNLVRK 294
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
459 LDFSSEELNFRFGLCSKLOMECPGKVLKEPMLVNSVNEQIKWTP-----VIRGFOEK 513
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 AEFSDAP--FAHEFGITINPANTEVKGRLVSLAPKLLYGGRTATALPNOGVWDMRG---K 349
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
514 QLVVVEPELCCAVFVNNETAGNCPLEENDVVVKFYTELIGGCKFRGIRIGANENGAOSI 573
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 QFHT---GIDVRVMAIACFAQQOHVKENDLRMFTNL-----QRI 386
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
574 MYDATKNEVAFYKNCNTLNTGIRPEIATEAKNMPERLPDKEQKVLMIILISKRQLNAYG 633
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
387 SNDAGMPIVG-----NPCFCYAVGEVQEPMPFKYLKQNYSGIQLVWVVLPGKTPVYA 439
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
634 FVKHYCDHTIGVANQHITSETVTKALASLRHEKSGKRIFYQIALKINAKLGGINGOLDWS 693
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
440 EVKRVGDTVLGIATQCVQAKNAIRT-----TPQTLNLCLENNVKLGGVNSIL--- 487
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
694 ETAETISPEEKERRKTMPLTMVYGVIDVTHPTSYSGIDYSIAAVASINPGGTIYRNMIVTQ 753
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
488 -----LPNVRPRIFNEPV-IPFGCDITHPPAGDSRKRPSIAAVVGSMDAHPRSRYAATVRVQ 541
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
754 BECRPGERAVAGRETDILEAKFVKLLREFAENNDNRAPAHIVVYRDGVSDSEMLRVSH 813
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
542 Q-----HRQELISDLTYMWRELLVQFYRTRFRK-PARIVYRDGVSEGGFFNVLQ 590
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
814 DELSLKSEVKQFMSERGEDPEPKYTFVIQKRHNTRLRRMEKDKPVNKNKDLTPAETD 873
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
591 YELRAIREAC--WMLER---GYQPGITTFIAVQKRHHTRLFAVDKKDQ--VGK----- 635
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
874 VAAVAVKQWEDMKESKETGI VNPSSGTTVDKLI VSKYKFPDFLASHHGVLGTSRPHYT 933
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
636 -----AYNIPPGETTVDVGITHPTTEFDYLCSHAGIQGTSRPSHYH 675
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
934 VMYDDKGMQSDEVYKMTYCLAFLSARCKPISLPVPVHYAHLSCAKELRYTYKEHYIG 993
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
676 VLWDDNNLTADLEQLLTQMCHTYVRCRNSVIPAPAYYAHILVAFRAR--YHLVDREHDSG 734
QY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
994 DYAAQP 998
Db      : : : :
735 EGSQP 739

```

Result	Query No.	Query			ID	Description
		Score	Match	Length		
c	1	5349	100.0	3229	AF180730	Caenorhab
	2	5001	93.5	39339	CEK08H10	Z83113 Caenorhabd
	3	2867	53.6	41648	CBG018K16	Z844520 Caenorhab

```
source
1. 3229
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/chromosomes="V"
/clone="yk296b10"
1. 3229
/gene="rde-1"
21..3083
/feature="rde-1"
/gene="rde-1"
/feature="rde-1"
/feature="similar to Caenorhabditis elegans K08H10.7"
/codon_start=1
/product="RNA interference promoting factor RDE-1"
/protein_id="AAF06159.1"
/db_xref="GI:6272678"
/translation="MSSNPFLEKGFYRHSLDPEMKWLARPTGKDGKPYEKKVLLV
NMFSSKTYDEYIEYEVKMTKEVLNRPKGFPPKTEIPIPDRAKLPWQHLREHK
QTFDLELVFDEKOTVSVCLNTVTSKMLVSEKVKVKKDEKDEKLEKILVTMI
LTVRKPHLNFRENPEDDEANRSYKFLKNVTKVYAPVFNSEIKVQPAKNFVYD
NNSILRVPEFDPNRPQESLEVAPRIENWFGIYIGIKELFDGEVLPALVDLFLYN
APMSLDYLLLIYVQSCNDVDRKDLKLMAGKMTINQAAPRIRQLLENLKKCA
EYDNMSRLTHERHITFDLCEENSLVYKVTGSDRGNKDYDTLFLKIEENKKFI
EFPFLPVKVSQAEYAVPMHELVHEKQYQRIIDLVMQDKFLKRAIRKPHDYKE
NTLKKLDFSEELNFFVERFGLCSKLMIECPGVKLEPMLVNSVEQIKMTIRIG
GQEQOLNVPEKELCCAVFVNENAGNPCLSENDVVKFYTELIGCKFRGIRIGANE
NRGAQSIMYDATNRYAFYKNCITLNTGIRFEIATEAKNMEERLPDKEQKLMFII
SRQLNAYGVKHYCDHTIGVANQHTISVTVKALASLHREKSGRIFYQIAKLNK
LGGINQEDMSIAISIEEKKRHTMPLTVMGIDVTHPTSISGIDYSIAAVASIN
PGGTYYNNIVQEECRPGERAVAGRETDILEAKFVKLREFAFENDRNRAPIHV
YRDGYSMLVSEKDELSLSEKQFMSERDSDPPKPYTFIVIQKHNTRILIRM
EKDKPVNKDLTPAETDVAAVYKOWEEDMKESKETGVNPSGGTVDKLIIVSKYKFD
PFLASHHGVLGTSRPGHTVMYVDKMSQDEVYKMTYGLAFLSARCSRPISLPVPVY
AHLSEKAKELRTYKEHYIGDYAQRTHHEHFLQTNVKYPGMSFA"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3229
Score: 5349.00 Matches: 1020
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0
US-10-645-746-3 (1-1020) x AFI80730 (1-3229)
QY 1 MetSerSerAsnPheProGluLeuGluLysGlyPheTyrArgHisSerLeuAspProGlu 20
Db 21 ATGTCCTCGAAATTTCCCGAAATGGAAAAAGGATTTTATCGTCAATCTCTCGATCCGGAG 80
QY 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLysVal 40
Db 81 ATGAAATGGCTTGCAGGCGCCCACTGGTAAATCGACGGCAAAATTTCTATGAGAAGAAAGTA 140
QY 41 LeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGlu 60
Db 141 CTTCCTTTGGTAAATTTGGTTCAAGTCTCCAGCAAAATTTACGATCGGGAATACTACGAG 200
QY 61 TyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLys 80
Db 201 TATGAAGTCAAAATGACAAAGGAAGTATTGAATAGAAAAACCAGGAAAAACCTTTCCCAAAA 260
QY 81 LysThrGluLysProLysPheArgAlaLysLeuPheTrpGluHisLeuArgHisGlu 100
Db 261 AAGACAGAAATTCCAATTCCTGAGTGTGCAAAACTCTTCTGGCAACATCTTCTGGCATGAG 320
QY 101 LysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAspThrValTyr 120
Db 321 AAGAAGCACACAGATTTTATCTCGAAGACTATGTTTTTGGATGAAAGGACACTGTTTAT 380
QY 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysValValLys 140
Db 381 AGTGTGTTGTCGACTGAACACTGTGCATCAATCAAAAAATGCTGGTTTCGGAGAAAGTAGTAAA 440
```

```
QY 141 LysAspSerGluLysLysAspGluLysLysAspLeuGluLysLysIleLeuTyrThrMetIle 160
Db 441 AAGGATTTGGAGAAAAAAGATGAAAAGGATTTGGAGAAAAAATAATCTTATACACAAATGATA 500
QY 161 LeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGluLysAspGlu 180
Db 501 CTTACCTATCGTAAAAAAATTTCCCTGAACCTTTAGTCGAGAAAAATCCCGAAAAAGACGAA 560
QY 181 GluAlaAsnArgSerTyrLysPheLysAsnValMetThrGlnLysValArgTyrAla 200
Db 561 GAAGCGAATCGGAGTTTACAAATTCCTGAAGAAATGTTATGACCCAGAAAGTTTCGTACGGC 620
QY 201 ProPheValAsnGluLysValGlnPheAlaLysAsnPheValTyrAspAsnAsn 220
Db 621 CTTTITTTGAAACGAGAGATTTAAAGTACAAATTCGGGAAAAAATTTTGTGTACGATATAAT 680
QY 221 SerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGlnSerLeuGlu 240
Db 681 TCAATTTCTCGAGTTTCTGAATCGTTTCAAGATCCAAACAGATTCGAACAATCATTAGAA 740
QY 241 ValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGlyLysLysGluLeuPheAsp 260
Db 741 GTAGCACCAGAAATCGAAGCATGGTTTGGAAATTTACATTTGGAATCAAGAAATTTGTTTCGAT 800
QY 261 GlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMet 280
Db 801 GGTGAACCTGTGCTCAATTTTGCATTTGTGCATTAACCTATTTCTACAAATCGACCGAAAAATG 860
QY 281 SerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAspAspValArg 300
Db 861 TCTCTCTCGATTATCTTCTCTTAATTTGTCCGACCCCGCATGCTGTAAACGATGATGATACGA 920
QY 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro 320
Db 921 AAAGATCTTAAACAAAACTGATGCGGGAATAATGACAAATCAGACAAACCGCGCGGCCA 980
QY 321 ArgIleArgGlnLeuLeuAsnLysLysCysAlaGluValTrpAspAsnGlu 340
Db 981 AGAATTCGACAAATTTATTTGGAATAATTTGAAGCTGGAATTCGAGAGAGTTTGGGATACGAA 1040
QY 341 MetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeuLysGluGluAsnSerLeu 360
Db 1041 ATGTCGAGATTGACAGAACGACATCTGACATTTCTAGATTTTGTGCGAGAAAACTCTCTT 1100
QY 361 ValTyrLysValThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThr 380
Db 1101 GTTTATAAGTCACCTGGTAAATCGGACAGAGAGAAATAATGCAAAAAAAGTAGTACGATCTACA 1160
QY 381 LeuPheLysIleTyrGluLysLysLysPheIleGluPheProHisLeuProLeuVal 400
Db 1161 TTGTTCAAAATCTATGAGGAAAAACAAAAGTTTCATTGAGTTTCCCCACCTACCACCTAGTC 1220
QY 401 LysValLysSerGlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGlu 420
Db 1221 AAAATTAAAGTGGAGCAAAAGATACGCTGTACCAATGGAATGGAACATCTTGAAGTTCTAG 1280
QY 421 LysProGlnArgTyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArg 440
Db 1281 AAGCCACAAAAGATACAAAGATCGAATTTGATCTGGTGTATGCAAGACAAAGTTTCTTAAAGCGA 1340
QY 441 AlaThrArgLysProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAsp 460
Db 1341 GCTACACGAAACCTCAGACTACAAAGAAAAATACCTTAAAAATTCGTGAAGAAATTTGGAT 1400
QY 461 PheSerSerGluGluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMet 480
Db 1401 TTCTCTTCTGAAGAGCTAAATTTTGTGTGAAGATTTGGATTTATGCTCCAAACTTCAGATG 1460
QY 481 IleGluCysProGlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGln 500
Db 1461 ATCGAATGTCCAGGAAGAGTTTGAAGAGCCCAATGCTTGTGATAGTGTAAATGAACAA 1520
QY 501 IleLysMetThrProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGlu 520
```



```

Db 1521 ATTAATATGACACAGTATTCGTGGATTTCCAGAAAACAAATGAAATGGTTCCCGAA 1580
Qy 521 LysGluLeuCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGlu 540
Db 1581 AAAGAACCTTGTCTGTCTGTTTGTAGTCAACGAAACAGCGGGAATCCATGCTTAGAA 1640
Qy 541 GluAsnAspValValLysPheThrGluLeuLeuGlyCysLysPheArgGlyLe 560
Db 1641 GAGAACGACGTTGTTAAGTTCTCACCGCACTAATGGTGGTTCCGAAGTCCCGGGAATA 1700
Qy 561 ArgLeuGlyAlaAsnGluAsnArgGlyValaGlnSerIleMetTyrAspAlaThrLysAsn 580
Db 1701 CGAATTGGTCCCAATGAAGACAGAGGAGCGCAATCTATTATGACGACGCGCAAGAAAT 1760
Qy 581 GluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyLeuGlyArgPheGluLeuAla 600
Db 1761 GAATATGCTCTTACAAAAATTTGACACATAATACCGGAATCCGGTAGATTGAATAGCC 1820
Qy 601 AlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLysValLeuMet 620
Db 1821 GCACAGACGAGGAAGATATGTTTGAACGCTCTCCCGATAAAGAACAAAAAGTCTTAATG 1880
Qy 621 PheLeuIleLeuSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHisTyrCysAsp 640
Db 1881 TTCATTATCATTTCCAAACGCAACTGAAATGCTTACGGTTTGTGAACCAATTATTGCGAT 1940
Qy 641 HisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAla 660
Db 1941 CACACCATCGGTAGCTAATCAGCATATTACTTCTGAAACAGTCACAAAAAGCTTTGGCA 2000
Qy 661 SerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsn 680
Db 2001 TCACTAAGGCACGAGAAAGATCAAAACGAAATTTCTATCAAAATGCAATGGAANAATCAAC 2060
Qy 681 AlaLysLeuGlyGlyIleAsnGlnGluLeuAspTyrSerGluIleAlaGluIleSerPro 700
Db 2061 GCGAATTAGAGGTATTACACGAGGCTTGACTGTCAGAAATGCGAANAATACCA 2120
Qy 701 GluGluLysGluArgArgLysThrMetProLeuThrMetTyrValGlyIleAspValThr 720
Db 2121 GAAGAAAAAGAAAGACGAAACAAATGCCAATTATGTTATGTTGGAATGATGTAAC 2180
Qy 721 HisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAlaSerIleAsn 740
Db 2181 CATCCAACTCCCTACAGTGGAAATTTGATTTATATAGCGGCTGTAGTAGCGAGTATCAAT 2240
Qy 741 ProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluCysArgProGlyGlu 760
Db 2241 CCAGGTGGAACTATCTATCGAATATGATTTGTGACTCAAGAGAATGTCTGCCGGTGAG 2300
Qy 761 ArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeu 780
Db 2301 CGTGCACTGCTCATGACGCGGAAACAGACAGATATTTTGGACGAAAGTTCGTGGAATTC 2360
Qy 781 LeuArgGluPheAlaGluAsnAsnAsnArgAlaProAlaHisIleValValTyrArg 800
Db 2361 CTCAGAGAAATTCAGAAACAAACGCAATCGAGCACCGCGCATATTGTAGTCTATCGA 2420
Qy 801 AspGlyValSerAspSerGluMetLeuArgValSerHisAspGluLeuArgSerLeuLys 820
Db 2421 GACGGATTAGCGATTCGGAGATGCTACGTGTTAGTATCATGATGAGCTTCGATCTTTAAAA 2480
Qy 821 SerGluValLysGlnPheMetSerGluArgAspGlyGluAspProGluProlLysTyrThr 840
Db 2481 AGCGAAGTAAACAAATTCATGTCGGAACGAGTGGAGAGATCCAGACGCGAGTAGACG 2540
Qy 841 PheIleValIleGlnLysArgHisAsnThrArgLeuLeuArgMetGluLysAspLys 860
Db 2541 TTCATTGTGATTCAGAAAAAGACAAATACACGATTCGTTGCTGGAAGATGGAAGATAAG 2600
Qy 861 ProValValAsnLysAspLeuThrProAlaGluThrAspValAlaValAlaValLys 880

```

```

Db 2601 CCAGTGTCAATAAAGATCTTACTCTCGTCAAAACAGATGCGCTGTTGCTGTGTTAAA 2660
Qy 881 GlnTrpGluAspMetLysGluSerLysGluThrGlyIleValAsnProSerSerGly 900
Db 2661 CAATGGGAGGAGGATATGAAAGAAAGCAAACTGGAATGTTGAACCCATCATCCGGA 2720
Qy 901 ThrThrValAspLysLeuIleValSerLysTyrLysPheAspPheLeuAlaSerHis 920
Db 2721 ACAACTGTGATAAACTTATCGTTTCGAAATACAAATTCGAATTTCTTCTGTCATCTCAT 2780
Qy 921 HisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMetTyrAspAspLysGly 940
Db 2781 CATGGTGTCTTGTGTACATCTCGTCCAGGACATTACACTGTTATGTATGACGATAAAGGA 2840
Qy 941 MetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCys 960
Db 2841 ATGAGCCAAAGTGAAGTCTATAAATGACCTACGGACTTGTCTTCTCTGCTAGATGT 2900
Qy 961 ArgLysProIleSerLeuProValProValHisTyrAlaHisLeuSerCysGluLysAla 980
Db 2901 CGAAAAACCCATCTCGTTGCTGCTTCCGGTTCAATTATGCTCATTTATCATGTGAAAAAGCG 2960
Qy 981 LysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAspTyrAlaGlnProArgThr 1000
Db 2961 AAAGAGCTTATCGAATTAAGAGAACATTACATCGGTGACTATGACAGCCACGGACT 3020
Qy 1001 ArgHisGluMetGluHisPheLeuGlnThrAsnValLysTyrProGlyMetSerPheAla 1020
Db 3021 CGACAGAAATGGAACATTTTCTCCAACTTAACGTGAAGTACCTCGAATGCTGTTCCGA 3080

```

RESULT 2

```

CEK08H10/c
LOCUS 39339 bp DNA linear INV 09-AUG-2005
DEFINITION Caenorhabditis elegans Cosmid K08H10, complete sequence.
ACCESSION Z83113
VERSION Z83113.1 GT:3217648
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 39339)
AUTHORS C. elegans Sequencing Consortium
CONSTRM Genome sequence of the nematode C. elegans: a platform for
TITLE Investigating biology
JOURNAL Science 282 (5396), 2012-2018 (1998)
PUBMED 9851916
REFERENCE 2 (bases 1 to 39339)
AUTHORS Gardner,A.E.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
worm@sanger.ac.uk

```

COMMENT

On Jun 13, 1998 this sequence version replaced gi:1695070.
Coding sequences below are predicted from computer analysis, using
Predictions from Genefinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
For a graphical representation of this sequence and its analysis
see: [http://www.wormbase.org/perl/ace/elegans/seq/sequence?](http://www.wormbase.org/perl/ace/elegans/seq/sequence?name=K08H10;class=Sequence)
name=K08H10;class=Sequence
This sequence is the entire insert of clone K08H10. The true left
end of clone T06B8 is at 32990 in this sequence. The true right end

of clone C27H6 is at 10790 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence Z81042. The end of this sequence (39233..39339) overlaps with the start of sequence Z73975.

FEATURES

Location/Qualifiers

1..39339
 /organism="Caenorhabditis elegans"
 /mol_type="genomic DNA"
 /strain="Bristol N2"
 /db_xref="taxon:62339"
 /chromosome="V"
 /clone="K08H10"

join(complement(638..744), complement(341..542), complement(105..294), complement(Z81042.1:24518..24775), complement(Z81042.1:24200..24465))
 /locus_tag="C27H6.8"
 join(complement(638..744), complement(341..542), complement(105..294), complement(Z81042.1:24518..24775), complement(Z81042.1:24200..24465))
 /locus_tag="C27H6.8"

/standard_name="C27H6.8"
 /note="contains similarity to Pfam domain PF03690 (Uncharacterised protein family (UPF0160))"
 /codon_start=1

/product="Hypothetical protein C27H6.8"
 /protein_id="CAB05549.1"
 /db_xref="GI:3924825"

/db_xref="InterPro:IPR003226"
 /db_xref="UniProtKB/Swiss-Prot:O17606"
 /translation="MTSLIGTSGKFKHCEAFKMLKQLPQKDSILTRDAAQLE KDIIVDVGIGIDHRSKORVDHQRGFTDMRTLEKLNPTKLSSAGLVAYHGREVIN QILGNEVSSMDLFFHRIYEQFVSIDALDNGISQYDGVPRYHSGNLSRGTQFNS HWNEPDAERQQAQMGPIGEFSRVKILANVWMPAREIIEAAVDKRFEDASGRI ILIENGCGPKWFFHDIIEVKNIADNNITILFSDTSNARVQAIPIVDKMSFENRM PLPAARWGLRDLDSKESGIPGGVFVHISGFTGCGNLTREGALAMARKALEIGENPVK KAKLGN"
 complement(join(1035..1256,1333..1815,1877..2485, 2530..2644,2685..3296,3343..3559,3607..3785,3840..3895, 4252,4295..4494,4550..4606))
 /gene="rde-1"

/locus_tag="K08H10.7"
 complement(join(1035..1256,1333..1815,1877..2485, 2530..2644,2685..3296,3343..3559,3607..3785,3840..3895, 3940..4252,4295..4494,4550..4606))
 /gene="rde-1"
 /locus_tag="K08H10.7"

/standard_name="K08H10.7"
 /note="C. elegans RDE-1 protein; contains similarity to Pfam domains PF02170 (PAZ domain), PF02171 (Piwi domain)"
 /codon_start=1

/product="Hypothetical protein K08H10.7"
 /protein_id="CAB05546.2"
 /db_xref="GI:14530521"
 /db_xref="InterPro:IPR003100"
 /db_xref="InterPro:IPR003165"

/db_xref="UniProtKB/TrEMBL:Q9XU82"
 /translation="MSSNPFLEKGFYRSLDPENKWLARPTCKDGFEYKKVLLLV NWFSSKIYDREYEVKMTKEVLNRPKPGFPFKTEIPIPDRAKLPWHLRHEKK QTFDILEDVFEKDTYVSVCLNTVTSMKVSEKVKDKSEKDKOLEKILYTM LTYRKHLNFRSDEPNKEDEANRSYKLVNMTQKRVYAPVNEEIKVQFAKNFYVD NNSILRVPSFDPNRFQOSLEVAPIEAWFGIYIGIKELPDGPEVLNPAIVDKLFYN APKMSLDYLLIIVDQSCNDVVRKDLTKLMAGKMTIRQARPRIFOLLNKLKCA EYWNEMSLTRHLTFDLCSNLSVYKVTQSKDRGNKAKYDITLKIYEENKFI EEPHLPLVKVSGAXEYVPMHEHLEVPHEKPKYKRNIDLVMDQDKFLKATRPKDYKE NTLKMLKELDFSSEELNPFVERGLCSKLMIECPGKLVKPEMLVNSVNEQKTPVIR GFOEQKLVNVPSEKELCAVFNVTAGNCPLEENDVVFYTBELIGCKPRGIRIGANE NRGQASIMYDANKYVAKNCTLTNTIGRFEIATAEAKMFERLPDKEQKVLMTIII SKQNLNAYGVFYCHDTIGVANQHTISETVTKALASIRHEKSKRIFVQIALKTNK LAGNIQELDWSIASISPEEKERRKTMPLMTVMGIDVTHPTSYSGIDVSIAAVASIN PGDTYRNMIIVTQECRPGERAVAHGRKTDILEAKFVKLLREFAEHNDNAPAHIVV YDGTGSDLSMLVDELSRLSKSVKQFMSERDGEDPEPKYTFYIQKRHNTRLRMR EKDKPVNKKDLTPAETDVAAVAKQWEDMKESKETGVNPSGGTVDKLIIVSKYKFD PFLASHHGVLGT"SRPGHYTVMYDDKMSQDEVYKMTYGLAFLSARCRKPIISLPVPVHY

gene

CDS

gene

CDS

AHLSCAKELRYTYKEHYIGDYAQRTRHMEHFLQTNVKKYPGMSFA"
 join(4899..4994,5051..5122,5167..5268,5318..5395, 5444..5584,5630..5698)
 /locus_tag="K08H10.9"
 join(4899..4994,5051..5122,5167..5268,5318..5395, 5444..5584,5630..5698)
 /locus_tag="K08H10.9"
 /standard_name="K08H10.9"
 /note="contains similarity to Pfam domain PF04051 (transport protein particle (TRAPP) component, Bet3)"
 /codon_start=1

/product="Hypothetical protein K08H10.9"
 /protein_id="CAB05547.1"
 /db_xref="GI:3924823"

/db_xref="InterPro:IPR007194"
 /db_xref="UniProtKB/TrEMBL:Q9XU81"
 /translation="MAFPFDILHLEIRTSERSEKRCRCAKEKFGTIIDNLNGPADIE KYESVQAFILHANAETKLESIGFRVGRQLVEKSKESQPLVTELEIVKFIKDFWSS VFGKVDNLRNTNHQGVVVDGREFTLRSFPEGQYVKESGYFLAPPCGLLRGALGL NIRAVTPTIESVPSKHIQVQPK"
 complement(join(6724..6781,6827..6943,6995..7242, 7294..7435,7684..8083,8127..8425,8700..8875))
 /locus_tag="K08H10.6"

complement(join(6724..6781,6827..6943,6995..7242, 7294..7435,7684..8083,8127..8425,8700..8875))
 /locus_tag="K08H10.6"

/standard_name="K08H10.6"
 /note="contains similarity to Pfam domain PF00083 (Sugar (and other) transporters)"
 /codon_start=1

/product="Hypothetical protein K08H10.6"
 /protein_id="CAB05545.1"

/db_xref="GI:3924821"
 /db_xref="GOA:Q9XU83"

/db_xref="InterPro:IPR005828"
 /db_xref="InterPro:IPR005829"

/db_xref="InterPro:IPR007114"

/db_xref="UniProtKB/TrEMBL:Q9XU83"
 /translation="MKVLPLERNLTPPHLTRLLATSVVATFAGGFHGYLLSAVNP LSELIQFIVDNLQNRYSLESSLQSLSIWSSAGLCFLGAMIGATYISILFQISQIPR NSLLAAIQLFSAFVGLAYSLRCELLPIRSLGIAFAIGIQAQVFLFEISPAK PRGNTNLSLGIENCAFLAATLTGTYVLTQWYTFIETIFCVLHLVNTTFH DSPTVLLSIGITSEGESAIKVYGETCHIKIMEDLRIDNGVKESLRQILKDKGCA QALSLAVINSPVSGIVAIPTFGTFLQITGFSPGSAVANSLCSPASIVSALIAA TAMDKIGRPILLITSLVALINIFMSILVELYDTSKOSFLAWPLFVLFTVFESI GIGPAAVFIGALAPPYISIKMQSYSTVQVQAGSFICPIIYLQLVESIGGFAPMLFIV FLTITAAFYMYLPETKQKSPVEIYGLIQ"
 join(17960..18067,18287..18427,18475..18790,19256..19589, 19640..19976,20048..20251)
 /gene="uda-1"

/locus_tag="K08H10.4"

join(17960..18067,18287..18427,18475..18790,19256..19589, 19640..19976,20048..20251)
 /gene="uda-1"

/locus_tag="K08H10.4"

/standard_name="K08H10.4"

/note="contains similarity to Pfam domain PF01150 (GDAI/CD39 (nucleoside phosphatase) family)"

/codon_start=1

/product="Hypothetical protein K08H10.4"
 /protein_id="CAB05544.1"

/db_xref="GI:3924820"

/db_xref="GOA:Q9XU84"

/db_xref="InterPro:IPR000407"

/db_xref="InterPro:IPR02034"

/db_xref="UniProtKB/TrEMBL:Q9XU84"

/translation="MLFPAPFSLISFSLSSVVTTKQYWCCHGVLNNQHTCRPFT IVIDAGSTGRLLHLVKFTHDPAIASHGMPFKVKEIFQEVKPGLSFSAKSPSSAASL EPLQARAKEPVHFWKTPITLKATAGRLLPAGMADDILRESVEERIPNSGFFAAPP DAVNWPQSDGCVSWFTNLILLETFLTDEPTVGHKPAHRSVAAPFLGSGSTQLYW PNNEAVFSEHVGYERDIDFFGHIIFLTSHFIGNGLIAARLNILQLETDNNEIESTHQL ITSQMEPYQJTEMEYALKVNEGNSSSFESCYTKTNKFNVESEIMHRELKQSPV YLFSYFFDRALNSGLVKGNEGKIELRQFKAAAEIACRREKTEIIDGSHWMPQCLDL

```

gene
    TTYISLLRRDQGQFEDNQPLVIAKKIKGMEVSWGQGLAFATANEPQLTGEGAIKTAALSE
    PNSTVVDQIFDLVYSGTNQVLSTSYNIIISV"
    join(24915..24953,26257..26580)
    /locus_tag="K08H10.3"
CDS
    join(24915..24953,26257..26580)
    /locus_tag="K08H10.3"
    /standard_name="K08H10.3"
    /note="contains similarity to Pteronotus parnellii Brcal
    (Fragment)"; TR:Q7YQI1"
    /codon_start=1
    /product="Hypothetical protein K08H10.3"
    /protein_id="CAB05542.1"
    /db_xref="GI:3924818"

```

Alignment Scores:

Pred. No.:	0	Length:	39339
Score:	5001.00	Matches:	1016
Percent Similarity:	85.2%	Conservative:	1
Best Local Similarity:	85.1%	Mismatches:	3
Query Match:	93.5%	Indels:	177
DB:	13	Gaps:	10

US-10-645-746-3 (1-1020) x CEK08H10 (1-39339)

Qy	1	MetSerSerAsn	PhePro	GluLeu	GlyPhe	TyrArg	HisSer	LeuAsp	Pro---	19
Db	4606	ATGTCCTCGAATTTT	CCCGAATTGGAAA	AAGGATTTTAT	CGTCATTCTCT	CGATCCGGTA	4547			

Qy 20 -----GluMetLy 22

Db
4546 TGATCAATTATTAGCAGCTATAAGATATATAAGTTTGATATTAATATATAGGAGATGAA 4487

QY 22 sTrpLeuAlaArgproThrGlyLysCysAspGlyLysphetYrGluLysLysValLeuLe 42

Db
4486 ATGGCTTGGAGGCCCACTGGTAAATGCGACGGCAAATCTATGAGAAGAAAGTACTTCT 4427

QY 42 uLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGI 62

Db 4426 TTTGGTAAATTGGTTCAAGTTC CAGCAAATTTACGATCGGAAATACGAGTATGA 4367

Qy 62 uValLysMetThrLysGluValLeuAsnArgLysProGlyLysPropheProLysLysTh 82

Db 4366 AGTGAAAAATGACAAAGGAAGTATTGAAATAGAAAAACGAGAAAACCTTTCCCAAAAAGAC 4307

Qy 82 rGluileProile-----ProAs 88

Db 4306 AGAAATTCCAATGTAAGTGCTTGTAATTAGTCAAACTAATTTTATTTTCAGTCCCGA 4247

QY 88 pArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysGlnThrAspPheLeLe 108

D_b 4246 TCGTGCAAACTCTTCTGGCAACATCTTCTGGCATGAGAAAGCAGACAGATTTTATTC 4187

QY 108 uGluAspTyrValPheAspGluLysAspThrValTyrSerValCysArgLeuAsnThrVa 128

Db 4186 CGAAGACTATGTTTTTTGATGAAAGGACACTGTTTATAGTGTTTGTGCTGAACACTGT 4127

Qy 128 IThrSerLysMetLeuValSerGluLysValValLysLysAspSerGluLysLysAspG1 148

Db 4126 CACATCAAAATGCTGGTTTCGGAGAAAGTAGTAAAAAGGATTTCGGAGAAAAAGATGA 4067

Qy 148 uLysAspLeuGluLysLysIleLeuTyrThrMetIleLeuThrTyrArgLysLysPheHi 168

Db 4066 AAAGGATTTCGGAGAAAAAATCTTATACACAATGATACTTACCTATCGTAAAAAATTTC 4007

Qy 168 sLeuAsnPheserArgGluAsnProGluLysAspGluAlaAsnArgSerTyrLysph 188

Db 4006 CCTGAAC TT TAGTCGAGAAAATCCGGAAAAGACGAAGCGAATCGGAGTTACAAATT 3947

Qy 188 eLeuLys-----AsnValMe 193

Db
3946 CCTGAA - GGTTTATGAAAAACACGCATTATAACAAACAAAATTAGCTTTCAGAAATGTTAT 3888

Qy 193 tThrGlnLysValArgTyrAlaProPheValAsnGluGluIleLysVal----- 209

Db	3887		GACCCAGAAAGTTCCGTACGCGCCTTTTGTGAACGAGGAGATTAAAGTGTGAGTTGCAAT	3828
Qy	210	-----	-----GlnPheAlaLysAenPh	215
Db	3827	AAATAATAATAATACCTCACTCACTTATATATATTTTAAAGACANATTCGCGAAAATTT	3768	
Qy	215	eValTyAspAenAenSerIleLeuArgValProGluSerPheHisAspProAenArgPh	235	
Db	3767	TGTGTACGATAATAATTTCAATTTCTCGAGTTCTCTGAATCGTTTTCACGATCCAAACAGATT	3708	
Qy	235	eGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyHleGlyIle	255	
Db	3707	CGAACAAATCATTAGAAGTAGCACCAAGATCGAAGCAGATGGTTTGGAAATTTACATTGGAAT	3648	
Qy	255	eLysGluLeuPheAspGlyGluProValLeuAenPheAlaIle	269	
Db	3647	CAAGAATTTGTCAGTGGTGACCTGTGCTCAATTTTGCAAG-TAAGTTTGAGAAACTGC	3589	
Qy	270	-----	-----ValAspLysLeuPheTyAsnAlaProLy	279
Db	3588	GATAAAAAATCATGTGATTTTGTGTGAAGTTGTGATAACTATTCTACAAATGCACCGAA	3529	
Qy	279	smetSerLeuLeuAspTyHleLeuLeuIleValAspProGlnSerCysAenAspAspVa	299	
Db	3528	AATGTCTTCTGTGATTTATCTCTCTTAATTTGCGACCCCGCAGTCGTGTGAACGATGATGT	3469	
Qy	299	laTyLysAspLeuLysThrLysLeuMetalGlyLysMetThrIleArgGlnAlaIleAar	319	
Db	3468	ACGAAAGATCTTAAACAAACTGTATGGCGGGAATGACAAATCAGACAAGCCCGCGC	3409	
Qy	319	gProArgIleArgGlnLeuLeuGlnAsnLeuLysCysAlaGluValTrpAspAes	339	
Db	3408	GCCAAAGATTCGACAAATTATTGGAAATTTGAAGCTGAAATGCGCAGAGTTTGGGATAA	3349	
Qy	339	nGluMet-----	-----SerArgL	344
Db	3348	CGAAATGTTAGTTTAAATTTATTCAAACAAATTAATATACAAATTTGATTTTCAGGTCCGAGAT	3289	
Qy	344	euThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAenSerLeuValTyLysV	364	
Db	3288	TGACAGAACGACATCTGACATTTCTAGATTTTGGCAGGAAACTCTCTGTTTATAAAG	3229	
Qy	364	alThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyAspThrThrLeuPheLysI	384	
Db	3228	TCAGTGTAAATCGGACAGAGGAAGAAATGCAAAAAGTACGATACTACATTTGTTCAAAA	3169	
Qy	384	leTyrgluGluAenLysLysPheIleGluPheProHisLeuProLeuValLysValLys	404	
Db	3168	TCTATGAGGAAAAACAAAAAGTTTCATTAGTTCCTCCACCTACCCACTAGTCAAAAGTTAAAA	3109	
Qy	404	erGlyAlaLysGluTyAlaValProMetGluHisLeuGluValHisGluLysProGlnA	424	
Db	3108	GTGAGCAAAAGAAATAGCTGTACCAATGGAACTTTGAAGTTTCATGAGAGCCACAAA	3049	
Qy	424	rgTyrgLysAenArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgL	444	
Db	3048	GATACAAGAAATCGAATTGATCTGTTGATCGACAGACAAAGTTTCTAAGCGAGCTACACGAA	2989	
Qy	444	ysProHisAspTyrgLysGluAenThrLeuLysMetLeuLysGluLeuAspPheSerSerg	464	
Db	2988	AACCTCAGCACTACAAAGAGAAATACCTTAAAAATGCTGAAAGAAATGGATTTCTCTCTGT	2929	
Qy	464	luGluLeuAenPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysP	484	
Db	2928	AAGAGCTAAATTTTGTGTGAAGAAATTTGGATTAATGCTTCCAACTTCAGATGATCGAATGTC	2869	
Qy	484	roGlyLysValLeuLysGluProMetLeuValAsnSerValAenGluGlnIleLysMetT	504	
Db	2868	CAGAAAGGTTTTGAAGAGCCAAATGCTTGTGAATAGTGAATGAACAAATTAATGA	2809	
Qy	504	hrProValIleArgGlyPheGlnGluLysGlnLeuAenValValProGluLysGluLeuL	524	

Db 2808 CACCACTGATTCGTGATTTTCAAGAAAAACAATTGAATGCTGGTTCCCGAAAAAGAACTTT 2749
 Qy 524 yscysAlaValPheValValAenGluThrAlaGlyAenProCysLeuGluGluAenAspV 544
 Db 2748 GCTGTGCTGCTTTTGTAGTCAACGAAACAGCGGAAATCCATGCTTTAGAGAGAACGACG 2689
 Qy 544 aIVal-----LysPheTyrThrGlu 550
 Db 2688 TTGTGTAGTGTTCCTACGTAGATTATTCGGAATAATATTTTCAGTAAGTTCTACCCGAA 2629
 Qy 551 LeuileGlyGlyCysLysPheArgGlyileAhrGlyleGlyAlaAenGluAenArgGlyAla 570
 Db 2628 CTAATTGGTGGTTGCAAGTTCCGTGGAATACGAAATGGTGGCAATGCAATGAAAAACAGAGAGCG 2569
 Qy 571 GlnSerileMetTyrAspAlaThrLysAenGlu----- 581
 Db 2568 CAATCTATTATGTACGACGCGAAGAAATGAATATGCGGTAAAGTTTCAGAAAAATTGAA 2509
 Qy 582 -----TyrAlaPheTyrLysAenCysThrLeuAenThrGlyileGly 595
 Db 2508 GTTTTAAATATCATATTACAG-TTCTACAAAAATTGTACACTAAATACCGAATCGGT 2450
 Qy 596 ArgPheGluileAlaAlaThrGluAlaLysAenMetPheGluArgLeuProAspLysGlu 615
 Db 2449 AGATTTGAAATAGCCGCAACAGAGCGAAGAATATGTTTGAACGTCTCCCGATAAGAA 2390
 Qy 616 GlnLysValLeuMetPheilellelleSerLysArgGlnLysAenAlaTyrGlyPheVal 635
 Db 2389 CAANAAGTCTTAATGTTCATTATCATTTCCAAACGACAACTGAATGCTTACCGTTTGTG 2330
 Qy 636 LysHisTyrCysAspHisThrileGlyValAlaAenGlnHisileThrSerGluThrVal 655
 Db 2329 AAACATTATGGATCACACCATCGGTGTAGTCTATCATCATATTACTTCTGNAACAGTC 2270
 Qy 656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgilePheTyrGlnile 675
 Db 2269 ACAAAAGCTTTGSCATCACTAAGGCACGAGAAAGGATCAAAACGAATTTTCTATCAAT 2210
 Qy 676 AlaLeuLysileAenAlaLysLeuGlyileAenGlnLysLeuAspTrpSerGluile 695
 Db 2209 GCATTTGAAATCAACGCGAAATTTAGGAGGTATTAACCGAGGCTTGACTGGTCAGAAAT 2150
 Qy 696 AlaGluileSerProGluLysGluArgArgLysThrMetProleuThrMetTyrVal 715
 Db 2149 GCAGAAATATCACCAGAGAAAGAAAGACGCGAACAATGCCATTATATGATGTT 2090
 Qy 716 GlyileAspValThrHisProThrSerTyrSerGlyileAspTyrSerileAlaVal 735
 Db 2089 GGAATTTGATGTAATCATCCAACTCCTACAGTGGAAATGATTTATTTATAGCGGCTGTA 2030
 Qy 736 ValAlaSerileAenProGlyGlyThrileTyrArgAenMetileValThrGlnGluGlu 755
 Db 2029 GTAGCGAGTATCAATCCAGGTGGAATATCTATCGAAATATGATGTGACTCAAGAAGAA 1970
 Qy 756 CysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspilleLeuGluAla 775
 Db 1969 TGTGTCGCGGTGAGGTGCGTGCATGACGCGGAAGAACAGATATTTTGGNAGCA 1910
 Qy 776 LysPheValLysLeuLeuArgGluPheAlaGlu----- 786
 Db 1909 AAGTTGCGTGAATTTGCTCAGAGAAATTCGCGAAGTGGAGTTGCTTTGAGTATTTAAAGAT 1850
 Qy 787 -----AenAenAspAenArgAlaProAlaH 795
 Db 1849 CTCTCGGATTTTAAATTTTTTGTAAACTTTTCAGAAACACGCAATTCGAGCACCGCGCA 1790
 Qy 795 sileValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspG 815
 Db 1789 TATTGTAGTCTATCGAGACGGAGTTAGCGATTCCGGAGATGCTACGTTAGTCATGATCA 1730
 Qy 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
 Db 1729 GCTTCGATCTTTAAAAAGCGAAGTAAACAATTTCTATGTCGGAACGGGATGGAGAAATCC 1670

Qy 835 oGluProLysTyrThrPheileValleGlnLysArgHisAenThrArgLeuLeuArgAr 855
 Db 1669 AGAGCCGAGATACAGTTTCAITGTGATTCAGAAAAGACACAATACAGATTGCTTGAAG 1610
 Qy 855 gMetGluLysAspLysProValValAasnLysAspLeuThrProAlaGluThrAspValAl 875
 Db 1609 AATGGAAAAAGATAAGCCAGTGGTCAATAAAGATCTTACTCTGCTGAAACAGATGTCG 1550
 Qy 875 aValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyileVa 895
 Db 1549 TGTCTGCTGTTTAAACAATGGAGAGGATATGAAGAAAGCAAGAAACATGGAAATGT 1490
 Qy 895 IaenProSerSerGlyThrThrValAspLysLeuileValSerLysThrLysPheAspPh 915
 Db 1489 GAACCCATCATCCGGAACAACATGTGGATAAATCTATCGTTTCGAAATACAAATTCGAT 1430
 Qy 915 ePheLeuAlaSerHisGlyValLysGlyThrSerArgProGlyHisTyrThrValMe 935
 Db 1429 TTTCTTGGCATCTCATCTGTTGTCCTTGGTACATCTCGTCCAGGACATTTACACTGTTAT 1370
 Qy 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyr----- 947
 Db 1369 GTATGACGATAAAGGAATGAGCCAAGATGAAGTCTATGTAAAGCTTTTGAATAGCAGTTA 1310
 Qy 948 -----LysMetT 950
 Db 1309 GCGATTTAGGATTTTGTAAATCCGCATATAGTTATATATAAAAAATGTTTCAGAAATGA 1250
 Qy 950 hrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProv 970
 Db 1249 CCTACGGAGTGTCTTCTCTCTGCTAGATGTGAAAAAGCCCATCTCGTTGCTGTTCCG 1190
 Qy 970 alHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluH 990
 Db 1189 TTCATTATGCTCATTTATCATGTGAAAAAGCCGAAAGAGCTTTATCGAACTTACAAGAAC 1130
 Qy 990 IsTyrileGlyAspTyrAlaGlnProArgThrArgHisGluMetGluHisPheLeuGlnT 1010
 Db 1129 ATTACATCGTGACTATGCACAGCCAGGACTCGACAGAAATGGNACATTTTCTCCAA 1070
 Qy 1010 hrAsnValLysTyrProGlyMetSerPheAla 1020
 Db 1069 CTAACGTGAAGTACCCTCGAATGTGTTCCGCA 1038

RESULT 3

CBRG18K16
 Caenorhabditis briggsae 41648 bp DNA linear INV 04-NOV-2000
 DEFINITION complete sequence.
 AC084520
 AC084520.1 GI:11094970
 HTG.
 Caenorhabditis briggsae
 Caenorhabditis briggsae
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditicoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 41648)
 Washington University Genome Sequencing Center.
 The C. briggsae Genome Sequencing Project
 Unpublished
 2 (bases 1 to 41648)
 Waterston, R.
 Direct Submission
 Submitted (04-NOV-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA
 e-mail: japieth@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone.

It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES

source
Location/Qualifiers
1. 41648
/organism="Caenorhabditis briggsae"
/mol_type="genomic DNA"
/strain="Gujarat G16"
/db_xref="taxon:6238"
/clone="G18K16"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 41648
Score: 2867.00 Matches: 638
Percent Similarity: 50.7% Conservative: 147
Best Local Similarity: 41.2% Mismatches: 203
Query Match: 53.6% Indels: 562
DB: 13 Gaps: 20

US-10-645-746-3 (1-1020) x CBRG18K16 (1-41648)

```

QY      6  ProGluLeuGluLys-----GlyPheTyrArgHisSerLeuAspProGlu----- 20
Db      7590 CCAAAATAGACAAAGAGCCAAATTAATACTACTCTCATCTGATCCAGT-GAGTCT 7648
QY      21 -----MetLysTrpLeuAla 25
Db      7649 TTAGACGCGCAAGAATCTAGAAAAAAATTCAGAGCATGAAGTGGTTATCA 7708
QY      26  AtqProThrClyLysCysAspGlyLysPheTyrGluLysLysValLeuValAsn 45
Db      7709 AGCCCAAGAGAAAAATGCGAAGGGAAGTATTATGTGCAAAAGTAAATTTGCTGGTCAAC 7768
QY      46  TrpPheLysPheSerLysLysLysLysLysLysLysLysLysLysLysLysLysLys 65
Db      7769 TGGTTCGATTTAACAGCAATATTACGATCGAGAGTATTATGAGTATTATTTGGATATG 7828
QY      66  ThrLysGlu-----ValLeuAsnArgLysProGlyLys-----Pro 77
Db      7829 TACAAAGAAAGAGATATATGGATAGAGAACTAAGTCGTTCAAAACAAAGCAGATGCCA 7888
QY      78  PhePro-----LysLysThrGluLeuProle-----Pro 87
Db      7889 TATCCATGTAAATGAAAAAGTAAACAATGAAGTCGTAATGTGCTTTTTCACGCCCG 7948
QY      88  AspArgAlaLysLeuPheTrpGlnHisLeuArgHisLysLysLysGlnThrAspPheIle 107
Db      7949 GAACGATCTGACGCTTCTGGAAGCACTTGAGATTCGAGCAACAAAGAAATCGTTCCAA 8008
QY      108 LeuGluAspTyrValPheAspGluLysAspThrValTyrSerValCysArgLeu---Asn 126
Db      8009 ATTGAAGACTATGTTTGTATGAGAAGGATACGGTTTACAGTATTACAAAGACCCCAAC 8068
QY      127 ThrValThrSerLysMetLeuValSerGluLysValValLysLysAspSerGluLysLys 146
Db      8069 GGAATGCTCTCGAAT-----CCTGATCTCTCAAAACCGA 8104
QY      147 AspGluLysAspLeuGluLysLysLysLysLysLysLysLysLysLysLysLysLys 166
Db      8105 GACATCAATATTAGTGCTTAA-----ATGCATTACAATGTATAA 8143
QY      167 PheHisLeuAsnPheSerArgGluAsnProGluLysAspGluLysLysLysLysLysLys 186
Db      8144 TTCAATCTATCTTTTAGTAGAAGATCCGACGAGGATGAAGCAAGCAAGCAAGTGT 8203
QY      187 LysPheLeuLys----- 190
Db      8204 AAATTTTGAAGGTTCTGTATTACTGTCTAAACAAACAATTTATTCGAAAGATTCA 8263
QY      191 -AsnValMetThrGlnLysValArgTyrAlaProPheVal----- 203
Db      8264 GAATGTTATGACACAAAAGTTGTTGTGACCGCTATGGCTTAATGAATCGGATTGTG 8323

```

```

QY      203 ----- 203
Db      8324 AGTCTGATTAAGTCTGAGACTTTAAAGAAAAATAACTCCGAAATTTTACCGGAATCGGGA 8383
QY      203 ----- 203
Db      8384 TGTTCGCAACATAATCTCTAAAAAATTTTCGACCAGACTACAGTACGTCAAGGCGCAGCGCA 8443
QY      203 ----- 203
Db      8444 AAAGTTTGCTTGGGGTTTCGAATGGNAATTACTTTCTAGTTTCAGGCTCTCCACAGACA 8503
QY      203 ----- 203
Db      8504 ATTCCAACTGCAAAATTAATTTTGTAAACTACAGAAAAAATTTGGCGCGATCACAATTCG 8563
QY      203 ----- 203
Db      8564 AGCTTTAAACAAGATTTTCAGAAATTTTATCTGATTTTGTGGAGTTCTCTAGATGATTA 8623
QY      203 ----- 203
Db      8624 AATGATTCCTTTAAAAAATAATTAAGTAGTAGTATAGATTCATCTATGTCCTTTTCGGGA 8683
QY      203 ----- 203
Db      8684 AAAAATGTCAAAAAATGTCATTTTGTAAAGTTTTTTACAGAAAAAATCTATCTCTCTT 8743
QY      203 ----- 203
Db      8744 TCAGTTTTTGACATTCCTAGGAGTCTATAGCTCAAAAAACCGCAATTAATTTCTCTACA 8803
QY      203 ----- 203
Db      8804 TTTTCTATGATGAATTTTAAATTCGTCTGATAAACCTTTTTTATTAATCAAAAAACAAC 8863
QY      204 -----AsnG 205
Db      8864 GCAGAAATGCGATTTGTGCAATTTTCATGAGTTTGTGNAATTTATAAAAAAAGTTTATCAG 8923
QY      205 LuGluLeuLysVal----- 209
Db      8924 ACGAATTTAAAAATTCATCTACAGAAAAATGTAAAGAAAAAATAATTTTTCGGTATTTTTCGAGCTA 8983
QY      209 ----- 209
Db      8984 TAGAACTCTAGAAATGTCAAAAACTGAAAGAGGAGATAGATTTTCTCTGTAATAAAACTT 9043
QY      209 ----- 209
Db      9044 TCAAAAATGACCAATTTTGACATTTTCTTCGGAAGAGGACCATACATGAATCTATAA 9103
QY      209 ----- 209
Db      9104 CCACTAATTAATTTTAAAGAAATCAATTTGAATCATCTTGGAACTCCACAAAAATCAGATAA 9163
QY      209 ----- 209
Db      9164 AAATTCGAATTAATGTTTGTAGACCTGCATTTGTGATCGCGCAAACTTTTCTGTAGA 9223
QY      209 ----- 209
Db      9224 TCAAAAATAAATTTGCAAGTTGCGAATTTGTTTGTGGGAGAGCGTGCAGTTTTTTTGTGCTT 9283
QY      209 ----- 209
Db      9284 TTGAAATCTTTTGTCAATTTTTCGGAATTTTCGAGCAATTCCTGTCCGCAAAATTAAT 9343
QY      209 ----- 209
Db      9344 TGAAAAATCGGTGGTAGCATATTTAAAGAAACCTGATGAATTTTGTGAGTGGCTGCTGCA 9403

```

```
QY 209 ----- 209
Db 9404 GCGAAACACCGTCTTACTAGTTTGGAAAAAATCTTTTTTTTCTAGGA 9463
QY 210 -----GlnPheAlaLysAsnPheValT 217
Db 9464 ATTTCATTTCCCGATTCCGAACAAGCGTTTTTTCAGCCAACTTGCCAAAATTTTCATCT 9523
QY 217 YrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluG 237
Db 9524 ACGATGAAACCTCCATCTCCGAGTTCCAGAAATCTTTTCCAGCACCTGATCGATTGAAC 9583
QY 237 InSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleValLeuG 257
Db 9584 ATTCTCGAAATTCGCGCACGATTAATCTGGTTCGGAATCTAATATCGCTGTTAAG 9643
QY 257 luLeuPheAspGlyGluProValLeuAsnPheAla ----- 268
Db 9644 AACTTTTCAGCGAGACCGGTTCTCAATTTTGGCAGTGAGAATTAATAATTTTAAGGCT 9703
QY 269 -----IleValAspLysLeuPheTyrsAsnAlaProLysMet 280
Db 9704 GAAATATTTAAACATTAATTTTCAGTTATTAAGCAATTTGTTTACAACTCCACAAATG 9763
QY 281 SerLeuLeuAspTyrrLeuLeuLeuIleValAspProGlnSerCysAsnAspValArg 300
Db 9764 TCTCTCTCGACTACTCTGTTCTGATTGGATTCGCAATCTCATATGGACGAAGTTGCA 9823
QY 301 LysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgPro 320
Db 9824 AGAAACCAATAAGACAGTTGAATTCGGGAATTTTGACAATTAAGCCACTATGTTTCAA 9883
QY 321 ArgIleArgGlnLeuLeuAsnLeuLysLeuLysCysAlaGluValTrpAsnGlu 340
Db 9884 AGCATTGATCAATTCGTTGCGAAATTTAAAGTTTAAAGCAGCATCGGTATGGAATCCTAA 9943
QY 341 -----MetSerArgLeu 344
Db 9944 AGGTATGAATCTACTATTGGAAGATGAGACGTTAATCAAAATTTTTCAGGTGAGGTTT 10003
QY 345 ThrGluArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrrLysVal 364
Db 10004 GCTGAACGTCATCTGACATTTATCAGATTGAGCAAGTATAATTTCCACGAAGAAATGATT 10063
QY 365 ThrGlyLysSerAspArgGlyArgAsnAlaLysLysTyrrAspThrThrLeuPheLysIle 384
Db 10064 CCAGTTCCAAAGAGGAAGACACAGAAATGCCCAATTCAGAGATTCCCATTTGTTCAAAT 10123
QY 385 TyrGluGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSer 404
Db 10124 TATGAGAGAACCGTAAATACATCCGATTTCTCGCTTACCATTTGGTGTGTGAAGAC 10183
QY 405 GlyAlaLysGluTyrrAlaValProMetGluHisLeuGluValHisGluLysProGlnArg 424
Db 10184 CGTAAATACAGATATTTCGGTCCCATCGGATTTCTAGAAAGTGTACGAAAAACCAACGC 10243
QY 425 TyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLys 444
Db 10244 TACAAAAATCGAATCCGATTTGCGATGCCAAGCAAAATTTGTGAATGCGACGACGCGAT 10303
QY 445 ProHisAspTyrrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGlu 464
Db 10304 CCTCATATGATTAAGAGGAGGCGCTTGAGATGTTGAAGATCTCGATTTTTCATCTGA 10363
QY 465 GluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysPro 484
Db 10364 TCGCTGAATCTTTGTTCAACCGTTTGGATTTCAGTACCGATCTGAAAATGATAGAGTGTGT 10423
QY 485 GlyLysValIleLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThr 504
Db 10424 GGAAGAGTCTGAAGAAACCATCCTCGTCAACAAAGATAATCAAAAGATCAGTATGACA 10483
QY 505 ProValIleArgGlyPheGlnGluLysGlnLeuAsnValValProGluLysGluLeuCys 524
```

```
Db 10484 CTTGGTGTAGGATTCCAGAGAAAGCGTTGAACGTGGTTCCAGAAAAAGAACTCTGT 10543
QY 525 CysAlaValPheValValAsnGlu---ThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
Db 10544 TGTGCCATTTTGTCTTCGCGAGCAAAAGACAGAGCCATGTTTGAAGAAGAAGAT 10603
QY 544 ValValLysPheTyrrThrGluLeuIleGlyCysLysPheArgGlyIleArgIleGly 563
Db 10604 GTGAGAGGTTCTACAAGACGCTCATCGATGTTGTGAATTCAGATCTATTGCAATCGGA 10663
QY 564 AlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrrAspAlaThrLysAsnGluTyrrAla 583
Db 10664 AAACATCAAAATAGTCAAGTTCTGTTCTTGTGCAAGATCCTGAAAGTGGGAAGTATGG 10723
QY 584 -----PheTyrrLysAsn 587
Db 10724 GTGAGTTGGTTTATCTCGAGTTAGTTACTGATATCCATTTATTTTTCAGTTCTACCCGAAC 10783
QY 588 CysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMet 607
Db 10784 GTGACGCTCAGTCGCGGCGGTACGCAACTTTAGAGCTGTGTCTAACGATCCGAAAGCAATG 10843
QY 608 PheGluArgLeuProAspLysGluGlnLysValLeuMetPheIleIleLysSerLysArg 627
Db 10844 TTTGAGAGACTTGTCTGACAGAATAACAAGATTCTACTGTTTCATCTGTTCTTTTCGAAACGT 10903
QY 628 GlnLeuAsnAlaTyrrGlyPheValLysHisTyrrCysAspHisThrIleGlyValAlaAsn 647
Db 10904 CGTTGGAATTTCTATGTTTCTGTAACAGTTCTGCGATGTTGAAAGTGTGTGTAGCAGT 10963
QY 648 GlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGly 667
Db 10964 CAACATGTTACTGCCACGTCGCGAAAGAGCTCTACAGGATATGTCTTCTATAAACCT 11023
QY 668 SerLysArgIlePheTyrrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsn 687
Db 11024 TCTAAACGAATTTCTATCAATCGCTTTAAAGATCAATGGAANAACCTAGGAGGTGTAAT 11083
QY 688 GlnGluLeuAspTrpSerGluIleAlaGluIleSerProGluGluLysGluArgArgLys 707
Db 11084 CAAGAACCTCGATTGTTGTCAGAAATCGAGAGATGACATGTTGAAAGAAAAAGAAAGAG 11143
QY 708 ThrMetProLeuThrMetTyrrValGlyIleAspValThrHisProThrSerTyrrSerGly 727
Db 11144 AATATGCCACTCCGAATGTATGTTGGATTGATGTCATCTATCCGACGAATGGAAGTGA 11203
QY 728 IleAspTyrrSerIleAlaAlaValValAlaSerIleAsnProGlyGlyThrIleTyrrArg 747
Db 11204 ATCGACTACTCGATAGTCTATTGTCGCAAGTATGATGATCCAGGAGGAACCGTTTATCGT 11263
QY 748 AsnMetIleValThrGlnGluLysCysArgProGlyGluArgAlaValAlaHisGlyArg 767
Db 11264 AATATGATTTGTAACCTCAAGAGAGAAATCGTCAGGTGAAACGTCGCGTAGCTCATGGAAG 11323
QY 768 GluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaGlu--- 786
Db 11324 GAGAGAACTGATATCTTGAACGGAAATTCGTGAAGCTTTTCAGAAATCTTTGACAGAGTA 11383
QY 787 -----AsnAs 788
Db 11384 TTAATCTTCTATTTTCAACGCGCAACATAGTTTGTGTTTCTATTGTTTTCAGAGAACAA 11443
QY 788 nAspAsnArgAlaProAlaHisIleValValTyrrArgAspGlyValSerAspSerGluMe 808
Db 11444 CGAAAACCGTATTTCCCTACATATGTTGTTTATCGAGACGGTGTGAGGACTCAGAAAT 11503
QY 808 tLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSe 828
Db 11504 GCTGGTGTGCTGATCATGATGAATTAAGGTCAATGAAGCGGAAGTTGAACGATTCCTGAA 11563
QY 828 rGluArgAspGlyGluAspProGluProLysTyrrThrPheIleValIleGlnLysArgHis 848
```


Pred. No.: 8 53e-57 Length: 4148
 Score: 582.00 Matches: 238
 Percent Similarity: 40.2% Conservatives: 172
 Best Local Similarity: 23.3% Mismatches: 370
 Query Match: 10.9% Indels: 243
 DB: 5 Gaps: 38

US-10-645-746-3 (1-1020) x BC063275 (1-4148)

Qy	21	MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrglu	37
Db	203	ATGGATGAA-GCGGACCTCGGAGCAGCTGCGGGCTTACCTGCCCCCTGCAG	261
Qy	38	-----LysLysValLeuLeu	42
Db	262	CAGGTGTTCCAGGCACCTCGCGGCTGGCATTTGGCAGTGTGGGAACCAATCAAGCTC	321
Qy	43	LeuValAsnTrpPheLysPheSerLysIleTyArgGluTyTyrglu	62
Db	322	CTGGCCNATTCTTGAGGTGGACATCCCTAAGATCGAGTGTACCACTACGAGGTGGAC	381
Qy	63	ValLysMetThrLysGluValLeuAsnArgLysProGlyLysPheProLysThr	82
Db	382	ATCAAGCCGATAAG-----TGTCCTCCGTAGAGTC	411
Qy	83	GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys	102
Db	412	AAACCGGAAGTGGTGGAA-----TACATGGTCCAGCATTTCAAGCCT	453
Qy	103	GlnThrAspPheIleLeuGluAspTy-----ValPheAspGluLysAspThrValTy	120
Db	454	CAG-----ATCTTGGTGTATCCCAAGCCTGTGTATGTATGGAAGAAGAACATTAC	504
Qy	121	SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal	138
Db	505	-----ACTGTACAGCACTGCCCATTTGGCAACGAACGGGTGCACTTT	546
Qy	139	---ValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysLysLys	157
Db	547	GAGGTGACAACTCCCTGGGAAGGAAGGAT---CGAATCTTTAAGGTCTCCATCAAGTGG	603
Qy	158	ThrMetIleLeuThrTyArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu	177
Db	604	CTAGCATTTGTAGCTGGCGAATGCTGCATGAGGCCCTGTGTCAGCGCCAGATCCCTGTT	663
Qy	178	LysAspGluAlaAsnArgSerTyLysPheLeuLysAsnValMetThrGlnLysVal	197
Db	664	CCCTTGAGTCTGTCAAGCCCTGTCATGTGGCCATGAGGCACCTGGCATCC-----ATG	717
Qy	198	ArgTyAlaPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTy	217
Db	718	AGGTACACCCCTGTG-----	732
Qy	218	AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln	237
Db	733	---GGCGGCTCTTCTTCTCAGCGCTGAGGGCTACTACACCCG-----	774
Qy	238	SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyIleGlyLysGlu	257
Db	775	-----CTGGGGGTGGGCGGAGGTCTGTTGCGCTTTCACCCAGTCTGTGGCCCT	825
Qy	258	LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyAsnAla	277
Db	826	GCATG---TGAAGATGATGTCTCAACATTTGATGTCTCAGCCACTGCTTTTATAGGCA	882
Qy	278	ProLysMetSerLeuLeuAspTyLeuLeuLeuIleValAspProGlnSerCysAsnAsp	297
Db	883	-----CAGCCAGTGATTGATGTGTGAGGTCTCGACATCAGGAACATAAATGAG	936
Qy	298	AspValArgLysAspLeuLysTyLysLeuMetAlaGlyLysMetThrIleArgGlnAla	317
Db	937	CAGCCCAAG-----CCCCCTCAGGACTCT	960

Qy	318	AlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys-----CysAla	334
Db	961	CAGCGCTTGGTTCCACAGAGATCAAGGCCCTGAAGGTGAAGTCAACCACTGTGGA	1020
Qy	335	GluValTrpAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr	350
Db	1021	CAGATGAAGAGGAAGTACCGCTGTGTAAATGTATTACCGCTCGCTGTAGCCATCAGACA	1080
Qy	351	PheLeuAspLeuCysGluGluAsnSerLeuValTyLysValThrGlyLysSerAspArg	370
Db	1081	TTCCCTTACAGCTGAGAGTGGACAGACTGTG-----	1113
Qy	371	GlyArgAsnAlaLysLysThrThrThrPheLysIleTyArgGluGluAsnLysLys	390
Db	1114	-----GAGTGCACAGTGGCACAGTATTTCAAGCAGAAATATAAC	1152
Qy	391	Phe-----IleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLysGluTy	409
Db	1153	CTTCAGCTCAAGTATCCCATCTGCTGCTTACAAGTTGGCCAGAACAAAAAGCATACC	1212
Qy	410	AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyLysAsnArgIle	429
Db	1213	TACCTTCCCTAGAGGTCTGTAAACATT---GTGGCTGGCGAGCGCTGTATTAAAAAGCTG	1269
Qy	430	AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyLys	449
Db	1270	ACCGACACACAGACCTCGACCATGATAAGGCCACAGCTAGATCCGCTCCAGACACAGAC	1329
Qy	450	GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal	469
Db	1330	GAGGAGATCAGTCGCTGATGAAGATGCCAGCTACAACTTAGATCCC-----TACATC	1383
Qy	470	GluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys	489
Db	1384	CAGGAATTTGGATCAAAAGTGAAGATGACATGACGAGGTGACAGGGGAGTGTGCTGCCG	1443
Qy	490	GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro-----	505
Db	1444	GGCCCATCTTGCAGTACGGCGCGGAACCGGGCCATTTGCCACACCCCAATCAGSGTGC	1503
Qy	506	---ValIleArgGlyPheGlnGluLysGlnLeu---AsnValValProGluLysGluLeu	523
Db	1504	TGGGACATCGCGGG-----AAACAGTTCTCAATGGATTTGAGATCAAAAGTCTGG	1554
Qy	524	CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp	543
Db	1555	GCATCGCTGCTTCCGCCCCCAAAA-----CAGTGTGAGAAAG-----	1596
Qy	544	ValValLysPheTyThrGluLeuIle-Gly-----GlyCysLysPh	557
Db	1597	GTGCTCAAGACTTCACAGCCAGCTGCGGAAGATTTCCAGAGTGTGCGGATGCTATC	1656
Qy	557	eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyAspAl	577
Db	1657	CAGGTCACCTTGTCTTCTGCAATATGACAGGGGGCA-GACAGCGTG-----	1704
Qy	577	aThrLysAsnGluTyAlaPheTyLysAsnCysThrLeuAsnThrGlyIleGlyArgPh	597
Db	1704	-----	1704
Qy	597	eGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnTy	617
Db	1705	-----GAGCTATGTTCCGGCATCTCTCAAGAACACCTACTCAGG	1742
Qy	617	sValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyGlyPheValLysHi	637
Db	1743	GCTGAGCTCATTTATTTCTCCAGGAAGACCGCGGTGTATGTGAGGTGAACG	1802
Qy	637	stTyCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLy	657
Db	1803	TGTCGAGATACACTCTTGGGAATGGCTACGAGGTGTGTGTCAGGTGAAGACGTGCTCA	1862
Qy	657	sAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyGlnIleAlaLe	677


```

Db      1863 GACC-----TCACCTCAGACTCTGCTCCAACTCTGCTCT 1895
Qy      677 uLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpSerGluIleAlaG1 697
Db      1896 CAAGATCAATGTCAAACTTGGTGCAATTAACAACATCCTA----- 1935
Qy      697 uIleSerProGluGluLysGluArgGlyThrMetProLeuThrMetTyrValGlyI1 717
Db      1936 -----GTCCACACACCGGCTCTGCCGTTTTTCAACAGCCAGTG---ATATTCTGGGAGC 1988
Qy      717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaValValAl 737
Db      1989 AGATTGTACACACCCCGACGAGGGATGGGAAAAACCTTCTATCACAGCAGTGGTAGG 2048
Qy      737 aSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluCysAr 757
Db      2049 CAGTATGATGATGCCACCCAGCGCATCTGTGCTACTGTGGGTACAG----- 2097
Qy      757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775
Db      2098 -----CGACCACCGCAAGAGATCATTTGAAGACTTTGTC 2129
Qy      775 aLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794
Db      2130 CTACATGTCGCTGAGCTCTCTCATCTCAATTTCTACAGTCCACCGTTTCAAG---CCTAC 2186
Qy      794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814
Db      2187 CCGCATCATCTTCTACCGAGATGGGTGCTGAAGGCCAGCTACCCAGATACCTCCACTA 2246
Qy      814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834
Db      2247 TGAGCTACTGCGCCATTCTGTGATCGCTGCATCAAACTG-----GAAAGGA 2291
Qy      834 pProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuAr 854
Db      2292 CTACCGAGCTGGATCACTTATTTATTTGGTGCGAAGCCATCACACCCGCTTTTC-- 2349
Qy      854 gArgMetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspVa 874
Db      2350 -TGTCGTGACAAATGAGCGAATTGGGAAGAGTGGTAACATCCAGCT----- 2397
Qy      874 lAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyI1 894
Db      2397 ----- 2397
Qy      894 eValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAs 914
Db      2398 -----GGGACCAAGTGGACCAACATCACCACCCATTGAGTTGA 2441
Qy      914 pPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrVa 934
Db      2442 CTTCTATCTGTGAGCCAGCCAGGCATCCAGGGCCAGCCAGCCAGCATCCCATCTATGT 2501
Qy      934 lMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAl 954
Db      2502 TCTTTGGGATGACAAACCGTTTTCACAGCATGAGCTCCAGATCCTGACGCTACCACTGTG 2561
Qy      954 aPheLeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHi 974
Db      2562 CCACACTTACGTACGATGCACGCTCTGTCTATCCAGCAGCCTGCTACTTACTGCCCG 2621
Qy      974 sLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHisTyrIleGlyAs 994
Db      2622 CCTGTGGCTTTCCGGGACGA---TACCACCTGCTGGTGGACAGGAGCATGACAGTGGAGA 2678
Qy      994 p 994
Db      2679 G 2679

```

RESULT 5
CQ727295

```

LOCUS      CQ727295              7478 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 13229 from Patent WO02068579.
ACCESSION  CQ727295
VERSION    CQ727295.1  GI:42292870
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS   Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE     Kits, such as nucleic acid arrays, comprising a majority of
          humanexons of transcripts, for detecting expression and other uses
          thereof
JOURNAL    Patent: WO 02068579-A 13229 06-SEP-2002;
          PE Corporation (NY) (US)
FEATURES   source
            1. 7478
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      7,87e-56      Length:      7478
Score:          577.00        Matches:    237
Percent Similarity: 40.2%      Conservative: 173
Best Local Similarity: 23.2%    Mismatches: 370
Query Match:    10.8%         Indels:     243
DB:             2             Gaps:        38
US-10-645-746-3 (1-1020) x CQ727295 (1-7478)
Qy      21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGlu----- 37
Db      209 ATGGGATGGAA-CGGGACCTCGGGAGCAGCTGCGGGCGCTTACTTGGCCCCCTCGCAG 267
Qy      38 -----LysLysValLeuLeu 42
Db      268 CAGGTGTTCCAGCACCCTCGCGCGCTGGCATTTGGCAGCTGTGGGAAACCAATCAAGCTC 327
Qy      43 LeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGluTyrTyrGluTyrGlu 62
Db      328 CTGGCCAAATTACTTTGAGGTGGACATCCCTTAAGATCGACGTGTACCACTACGAGGTGGAC 387
Qy      63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82
Db      388 ATCAAGCCGGATTAAG-----TGTCCTCCGTAGAGTC 417
Qy      83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
Db      418 AACCGGGAAGTGGTGAA-----TACATGGTCCAGCATTTCAAGCCT 459
Qy      103 GlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLysAspThrValTyr 120
Db      460 CAG-----ATCTTTGGTGTATCGCAAGCCCTGTGTATGATCGAAAGAAACAATTTAC 510
Qy      121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal----- 138
Db      511 -----ACTGTCAAGCATCCCATTTGGCAACGAGGTCGAGCTTT 552
Qy      139 ---ValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysIleLeuTyr 157
Db      553 GAGGTGACAATCCCTGGGGAAGGAAGAT---CGAATCTTTAAGGTCTCTCATCAAGTGG 609
Qy      158 ThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
Db      610 CTAGCCATTGTGAGCTGGCGGAATGTGTCATGAGGCCCTGTGTCTAGCGGCCAGATCCCTGTT 669
Qy      178 LysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGlnLysVal 197
Db      670 CCGTTGGAGTCTGTGCNAAGCCCTCGGATGTGGCCATGAGGCACCTGGCATCC-----ATG 723

```

QY 198 ArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsnPheValTyr 217
DB AGGTACACCCCTGTG----- 738

QY 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
DB GGGCGCTCTCTTCTTCACCGCTGAGGGCTACTACACCCG----- 780

QY 238 SerLeuGluValAlaProArgIleGluAlaTyrPheGlyIleTyrIleGlyIleLysGlu 257
DB CTGGGGGGTGGCGCGAGGCTGTGTCGGCTTTCACCAAGTCTGTGCGCCCT 831

QY 258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyrAsnAla 277
DB TGGAAAGATGATGCTCAACATTTGATGCTCAGCCACTGCCTTTTATAGGCA 888

QY 278 ProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSerCysAsnAsp 297
DB CAGCCAGTGATTCAGTTTCATGTGTGAGGTGCTGGACATCAGGAACATAGATGAG 942

QY 298 AspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIleArgGlnAla 317
DB CAGCCCAAG-----CCCTCAGCGACTCT 966

QY 318 AlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys-----CysAla 334
DB CAGCGCGTTTCCTCACCAGAGAGATCAAGGGCCTGAAGGTGGAAAGTCAACCCACTGTGGA 1026

QY 335 GluValTyrAspAsn-----GluMetSerArgLeuThrGluArgHisLeuThr 350
DB CAGATGAAGAGAAAGTACCGCTGTGTAATGTTTACCGCTGCGCCCTGTAGCCATCAGACA 1086

QY 351 PheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAspArg 370
DB TTCCCTTACAGCTGGAGAGTGACACTGTG----- 1119

QY 371 GlyArgAsnAlaLysLysTyrAspThrLeuPheLysIleTyrGluGluAsnLysLys 390
DB -----GAGTGCACAGTGGCACAGTATTTTCAAGCAGAAATATAAC 1158

QY 391 Phe-----IleGluPheProHisLeuValLysValLysSerGlyAlaLysGluTyr 409
DB CTTACAGTCAAGTATCCCATCTGCTCCCTCCTCACAAGTTGGCCAGGAACAAAGATACC 1218

QY 410 AlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsnArgIle 429
DB TACCTTCCCTAGAGTCTGTAAACATT---GTGGCTGGCGAGCGCTGTATTAAAAAGCTG 1275

QY 430 AspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLysProHisAspTyrLys 449
DB ACCGCAACACAGCTCGACCATGATAAAGGCCACAGCTAGATCCGCTCCAGACAGACAG 1335

QY 450 GluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLeuAsnPheVal 469
DB GAGGAGATCAGTCGCTGTAGAGATGCGACACTACACTTAGATCCC-----TACATC 1389

QY 470 GluArgPheGlyLysSerLysLeuGlnMetIleGluCysProGlyLysValLeuLys 489
DB CAGGAATTTGGGATCAAAAGTGAAGATGACATGACGAGGTGACAGGGCGAGTGTGCGG 1449

QY 490 GluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro----- 505
DB GCGCCCATCTTCAGTACGCGCGCGGAACCGGGCCATTGGCCACACCCCAATCAGGGTGT 1509

QY 506 ---ValIleArgGlyPheGlnGluLysGlnLeu---AsnValValProGluLysGluLeu 523
DB TGGGACATCGGGGG-----AAACAGTTCTACATGGGATTCAGATCAAGTCTGG 1560

QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluAsnAsp 543
DB GCCATCGCTGCTTGCACCCCAAAAA-----CAGTGTGAGAAAGAG----- 1602

QY 544 ValValLysPheTyrThrGluLeuIle-Gly-----GlyCysLysPhe 557
DB GTGCTCAAGAACTTCACAGACAGCTGCGGAAGATTTCCAAGATGCGGGGATGCTTATC 1662

QY 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIleMetTyrAspAl 577
DB CAGGGTCAACCTTGTCTGCAATATGACAGAGGGGCA-GACAGGGT----- 1710

QY 577 aThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleGlyArgPhe 597
DB ----- 1710

QY 597 eGluIleAlaIaThrGluAlaLysAsnMetPheGluArgLeuProAspLysGluGlnLys 617
DB -----GAGCCTATGTTCCGGCATCTCAAGAACACCTTACTCAGG 1748

QY 617 eValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheValLysHis 637
DB ----- 1749

QY 637 eTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrValThrLys 657
DB TGTGGAGATACACTCTTGGGAATGCTACGCACTGTGTGCAAGTGAAGAAGCTGTGTCAA 1868

QY 657 eAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnIleAlaLe 677
DB GACC-----TCACCTCAGACTCTGTCTCAACCTCTGCGCT 1901

QY 677 uLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTyrSerGluIleAlaGln 697
DB CAAGATCAATGTCAAACTTGTGGCATTAACAACATCTTA----- 1941

QY 697 uLysSerProGluGluLysGluArgLysThrMetProLeuThrMetTyrValGlyIle 717
DB -----GTCCACACACCGCTCTGCGTTTTTCAACAGCCAGTGT---ATATCTCTGGAGC 1994

QY 717 eAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAlaValAla 737
DB ----- 1995

QY 737 aSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGluCysArg 757
DB ----- 2053

QY 757 gProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGlu-----Al 775

DB -----CGACCCAGCAAGAGATCATTTGAAGACTTGTTC 2135

QY 775 aLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl 794

DB CTACATGTGTGGTGAAGTCTCTCATCCATTTCTACAGTCCACCCCGTTTCAAG---CCTAC 2192

QY 794 aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs 814

DB CGGCATCATCTTCTACCGAGATGGGTGCTGTAAGGCCAGCTACCCAGATATCTCCACTA 2252

QY 814 pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs 834

DB TGAGTACTGCGCATTCGTGATGCTGTCATCAAACTG-----GAAAGGA 2297

QY 834 pProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnThrArgLeuLeuArg 854

DB CTACCGCTGGGATCACTTATTTGTGTGCGAAGAACGCCATCACCCCGCTTTC--- 2355

QY 854 gArgMetGluLysAspLysProValAlaAsnLysAspLeuThrProAlaGluThrAspVa 874

DB -TGTCTCACAAGATGACGGAATTTGGGAAGAGTGTAACTCCAGCT----- 2403

QY 874 lAlaValAlaIaValLysGlnTyrGluGluAspMetLysGluSerLysGluThrGlyIle 894

DB ----- 2403

QY 894 eValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAs 914

Db 2404 -----GGGACCAAGTGGACCAACATACCCACCCATTTGATTTGA 2447
 Qy 914 pPheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisThrVa 934
 Db 2448 CTTCTATCTGTGCAGCCACGAGCATCCAGGCGACCGACCATCTACTATGT 2507
 Qy 934 lMetTyAspAspGlyMetSerGlnAspGluValTyLysMetThrTyGlyLeuAl 954
 Db 2508 TCITGGGATGACAAACCGTTTACAGAGATGAGCTCCAGATCCGAGTACAGCTGTG 2567
 Qy 954 aPheLeuSerAlaArgCysArgLysProIleSerLeuProValHisTyAlaHi 974
 Db 2568 CCACACTTACGTAGTGCACAGCTCTGTCTATCCAGCACCTGCTACTATGCCG 2627
 Qy 974 sLeuSerCysGluLysAlaLysGluLeuTyArgThrTyLysGluHisTyLys 994
 Db 2628 CCTGGTGGCTTTCGGGACGA---TACCACCTGGTGACAGAGCATGACAGTGGAGA 2684
 Qy 994 p 994
 Db 2685 G 2685

RESULT 6
 AF093097
 LOCUS Homo sapiens putative RNA-binding protein Q99 mRNA, complete cds.
 AF093097
 VERSION AF093097.1 GI:6002622
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 7478)
 Koesters, R., Adams, V., Betts, D., Moos, R., Schmid, M., Siermann, A.,
 Doeberitz, M., Weitz, S., Lichter, P., Heitz, P.U., von Knebel
 Hassem, S., and Briner, J.
 Human eukaryotic initiation factor EIF2C1 gene: cDNA sequence,
 genomic organization, localization to chromosome bands 1p34-p35,
 and expression
 Genomics 61 (2), 210-218 (1999)
 10534406
 2 (bases 1 to 7478)
 Koesters, R., Briner, J., Moos, R., Schmid, M., Doeberitz, M.V.K. and
 Betts, D.
 Direct Submission
 Submitted (18-SEP-1998) Surgery, University Hospital, INF 110,
 Heidelberg 69120, Germany
 Location/Qualifiers
 1. 7478
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosomes="1"
 /map="1p34-p35"
 214..2787
 /notes="similar to Arabidopsis thaliana AGO1"
 /codon_start=1
 /product="putative RNA-binding protein Q99"
 /protein_id="AA00068.1"
 /db_xref="GI:6002623"

CDS
 /translation="MSAGPSGAAGAYLPPLQQVFOAPRRPGITVGKPKILANYFE
 VDIKIDVHYEVDIKDPCPRVRNREVYVQHPKQIFGDRKVPYDGKNIYTVT
 ALPIGNREVDFVTIPEGKDRIFKVISIKWLAIVSRWLHVALVSGQIIPVLESVAL
 DVAMHLSMRVTPVGRSFPSPGYTHPLGGREWFPGHQSVPAPKWMNLIDVS
 ATAFYKQVIFSMCEVLDIRNIDEQPELTSQRVRFTGKINGLVETHCGQMRK
 YRVNTRPASHQRTPLLESQGVTECTVAQYFKOKYNLQKLYPLHPLCLQVQEQKH
 TYLPLEVCNIPAGQRCIKKLTQNTSTMIKATARSAPDQEEISRLMKQVSNLDPYI
 QBFIKVDMDTEVTRVLPAFLQYGGNRRAIATPNQGWDMRGKQFYNGIEIKWA
 IACFAPKQCRREVLNFTDRLKISKDGMPIQGPCFCYKAGQADSPEMPHILKN
 TYSGLQLIIVLPKTPVYAEVKRVGDTLLGNATQCQVQVKNVKTSPQTLNLCLKN

ORIGIN
 Alignment Scores: 7.87e-56 Length: 7478
 Pred. No.: 577.00 Matches: 237
 Score: 577.00 Matches: 237
 Percent Similarity: 40.2% Conservative: 173
 Best Local Similarity: 23.2% Mismatches: 370
 Query Match: 10.8% Gaps: 243
 DB: 5
 US-10-645-746-3 (1-1020) x AF093097 (1-7478)
 Qy 21 MetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyGlu----- 37
 Db 209 ATGGGATGGAA-GCGGACCCCTCGGAGCAGCTGCGGCGCTTACCTGCCCTCCCTGCAG 267
 Qy 38 -----LysLysValLeuLeu 42
 Db 268 CAGGTGTTCCAGCACCTCGCCGCTGGCATTTGGCAGCTGGGGAAACCAATCAAGCTC 327
 Qy 43 LeuValAsnTrpPheLysPheSerLysIleTyAspArgGluTyTyGluTyGlu 62
 Db 328 CTGGCCAATTACTTTGAGGTGGACATCCCTAAGATCGACGTGTACCACCTACGAGGTGCAC 387
 Qy 63 ValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPheProLysLysThr 82
 Db 388 ATCAAGCCGGATAGG-----TGTCCTCCCGTAGATGC 417
 Qy 83 GluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArgHisGluLysLys 102
 Db 418 ACCCGGAAGTGTGGAA-----TACATGGTCCAGCATTTCAAGCCT 459
 Qy 103 GlnThrAspPheIleLeuGluAspTy-----ValPheAspGlyLysAspThrValTy 120
 Db 460 CAG-----ATCTTTGTGTATCGCAAGCCTGTGTATGATGATGAAAGAAAGAACATTTC 510
 Qy 121 SerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLysVal----- 138
 Db 511 -----ACTGTACAGCAGCTGCCCATTTGGCAACGACGCGTGCAGCTTT 552
 Qy 139 ---ValLysLysAspSerGluLysLysAspGluLysAspLeuLysLysIleLeuTy 157
 Db 553 GAGGTGACAACTCCCTGGGAAGGAAGGAT---CGAATCTTTAAGTCTCCATCAAGTGG 609
 Qy 158 ThrMetIleLeuThrTyArgLysLysPheHisLeuAsnPheSerArgGluAsnProGlu 177
 Db 610 CTAGCCATTGTGAGCTGGCGAATGCTGCATGAGGCCCTGTGTACGCGCCAGATCCCTGTT 669
 Qy 178 LysAspGluGluAlaAsnArgSerTyLysPheLeuLysAsnValMetThrGlnLysVal 197
 Db 670 CCTTGGAGTCTGTGCCAACCCCTGGATGTGGCATGAGGCACCTGGCATCC-----ANG 723
 Qy 198 ArgTyAlaProPheValAsnGluGluIleLysValIleGlnPheAlaLysAsnPheValTy 217
 Db 724 AGGTACACCCCTGTG----- 738
 Qy 218 AspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArgPheGluGln 237
 Db 739 ---GGCCGCTCTCTTCTTCTCACCGCTGAGGGCTACTACACCCG----- 780
 Qy 238 SerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyLysGlyLysGlu 257
 Db 781 -----CTGGGGGTGGGCGCGAAGTCTGGTTCGGCTTCCACGAGTGTGTGCGCCCT 831
 Qy 258 LeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPheTyAsnAla 277
 Db 832 GCATG---TGGAAAGATGATGCTCAACATTGATGTCTCAGCCACTGCTCTTTTATAAGGCA 888

Db	2628	CCTGGTGGCTTTCGGGCACGA	---	TACCACCTGGTGACAAGAGCATGACAGTGGAGA	2684
Qy	994	G	994	G	994
Db	2685	G	2685	G	2685
RESULT 7					
LOCUS	AR439853				
DEFINITION	Sequence 105 from patent US 6664446.				
ACCESSION	AR439853				
VERSION	AR439853.1	GI:42665818			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 2910)				
AUTHORS	Heard,J., Brown,P., Riechmann,J.L., Keddie,J., Pineda,O., Adam,L., Samaha,R., Zhang,J., Yu,G.-L., Ratcliffe,O., Pilgrim,M., Jiang,C.-Z. and Reuber,L.				
TITLE	Transgenic plants comprising polynucleotides encoding transcription factors that confer disease tolerance				
JOURNAL	Patent: US 6664446-A 105 16-DEC-2003;				
FEATURES	Mendel Biotechnology, Inc.; Hayward, CA				
source	location/Qualifiers				
	1..2910				
ORIGIN	/organism="unknown"				
	/mol_type="genomic DNA"				
Alignment Scores:					
Pred. NO.:	2,95e-56	Length:	2910		
Score:	575.50	Matches:	237		
Percent Similarity:	39.0%	Conservative:	159		
Best Local Similarity:	23.3%	Mismatches:	337		
Query Match:	10.8%	Indels:	283		
DB:	2	Gaps:	41		
US-10-645-746-3 (1-1020) x AR439853 (1-2910)					
Qy	38	LysLysValLeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyrAspArgGlu	57		
Db	472	AAGAAAGTCATGGTTCGTGGCAAT-----CATTTCTGTTCAAGTTCGTATCGTGAT	525		
Qy	58	TyrTyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro	77		
Db	526	CTCTACCATTTACGATGTTTCGATCAATCCTGAGGTTATATCAAG-----	570		
Qy	78	PheProLysLysThrGluIlePheProLeuProAspArgAlaLysLeuPheTrpGlnHisLeu	97		
Db	570	-----	570		
Qy	98	ArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyrValPheAspGluLysAsp	117		
Db	570	-----	570		
Qy	118	ThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGluLys	137		
Db	571	-----ACAGTGAACAGAAACGTGATGAAACTTCTGTTAAAGAAAT---	609		
Qy	138	ValValLysLysAspSerGlu-----LysLysAspGluLysAspLeuGluLysLys	154		
Db	610	-----TATTAAGAACTCTCACTTGGGAGGGAAGTCCACGCGTATGAT---	660		
Qy	155	IleLeuTyrThrMet-----IleLeuThrTyrArgLysLysPheHisLeuAsnPheSer	172		
Db	661	AGCCTCTACTGCTGCTCCATACCTTTTGACTCGAAGAGTTGTTGTGATCTGGCG	720		
Qy	173	ArgGluAsnProGluLysAspGluGluAlaAsnArgSerTyrLysPhe-----LeuLysAsn	191		
Db	721	GAGAAAGAGCTGACCGGTTCTCTCGGGAAGGACAGACCGTTTAAAGTTGCTGTAAGAAT	780		


```
QY 613 pLysGluGlnLysValLeuMetPheIleIleIleSerLysArgGlnLeuAsnAlaTyrGl 633
Db 1524 TACCTACTCAGGACTGCCTTATTCGTCTGCGCCGGAAGACGCCAGTGTATGC 1583
QY 633 yPheValLysHisTyrCyAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGl 653
Db 1584 TGAAGTGAACGTGTCGGAGATACACTCTTGGGATGGCAACACACTGTGTGCAGGTGA 1643
QY 653 uThrValThrLysLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTy 673
Db 1644 GAATGCGGTCAAGACG-----TCACCGACGACTCTGTC 1676
QY 673 rGlnIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpse 693
Db 1677 CAACCTCTCCCTCAAGATCAATGTCAAACTCGGTGGCATTAACAACATCCTTA----- 1728
QY 693 rGluIleAlaGluIleSerProGluGluLysGluArgLysThrMetProLeuThrMe 713
Db 1729 -----GTCCACACCGCTCTGCTGCTTTTTCACACCGAGTG---AT 1769
QY 713 tTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAl 733
Db 1770 TTTCTGGGAGCAGATGTACACATCCCGCAGCTGGGATGGGAAGAAACCATCTATCAC 1829
QY 733 alaValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGl 753
Db 1830 AGCAGTGGTAGGCAGCATGGACGCACACCCCGCATCTGTGCCACCGTCGTGTGCA 1889
QY 753 nGluGluCyAspProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLe 773
Db 1890 G-----AlaLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAs 1910
QY 773 uGlu-----AlaLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAspAs 790
Db 1911 TGAAGACTTGTCTATATGTGCGCAGCTGTCTTATCCAGTCTTACAGTCCACCCGCTT 1970
QY 790 nArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuAr 810
Db 1971 CAAG---CCACCCCGCATCTTCTTACCGCAGCGCTCCCGGAGGCGAGTCCGCCA 2027
QY 810 gValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluAr 830
Db 2028 GATTTCTTACTATGACTGCTTGCCTTCCAGTATGCATGCATCAAACTG----- 2076
QY 830 gAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAsnTh 850
Db 2077 ---GAAAAAGACTACCGCTGGGATCATATATTGTGTGTCAGAGCGCCATCACAC 2132
QY 850 rArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrProAl 870
Db 2133 CGCCTGTTC---TGTGCTGACAAAGATGAGCGGATGGAGAGTGGTAACATCCACG 2189
QY 870 aGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLysGluSerLy 890
Db 2190 A----- 2190
QY 890 sGluThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSerLy 910
Db 2191 -----GGGACCACTGTGGACACCAACATCACCCACCC 2222
QY 910 sTyrLysPheAspPheLeuAlaSerHisGlyValLeuGlyThrSerArgProGl 930
Db 2223 ATTGTAGTTCGACTTCTATCTGTGCGACCATTCAGGCGATTCGAGGTAACCGGACCATC 2282
QY 930 yHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetTh 950
Db 2283 CCATTATTATGCTCTTGGATGACACCCGTTTCACGACGATGAATCCAGATCTTGAC 2342
QY 950 rTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValProVa 970
Db 2343 ATACCAAGCTGTGCCACACTTACGTACGATGCACACGTTCTGTCTATCCCGACACCTGC 2402
```

```
QY 970 lHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHi 990
Db 2403 CTACTATGCTCCCTTGTGGCTTCGGGCACGA---TACCACCTGGTGCAGACGACACA 2459
QY 990 sTyrIleGlyAsp 994
Db 2460 TGACAGTGGAGAG 2472

RESULT 9
BC077863 3298 bp mRNA linear VRT 12-OCT-2004
Xenopus laevis eukaryotic translation initiation factor 2C, 1, mRNA
(cDNA clone MGC:80582 IMAGE:5440319), complete cds.

BC077863
BC077863.1 GI:50418262
MGC.
Xenopus laevis (African clawed frog)
Xenopus laevis
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus;
1 (bases 1 to 3298)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 3298)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
Mckernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 3298)
Klein, S. and Gerhard, D.S.
Direct Submission
Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Anura Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angeliqne Schnerch, Ursula Skalska,
```


Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 154 Row: c Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

```

Location/Qualifiers
1..3298
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:80582 IMAGE:5440319"
/tissue_type="Embryo, stage 31/32, Xenopus"
/clone_lib="NICHG_XGC_Emb4"
/lab_host="DH108"
/note="Vector: pCMV-SPORT6"

gene
1..3298
/genes="eif2c1-prov"
/note="synonym: MGC80582"
/db_xref="GeneID:446823"
263..2851
/genes="eif2c1-prov"
/codon_start=1
/product="eif2c1-prov protein"
/protein_id="AAH7863.1"
/db_xref="GI:50418263"
/db_xref="GeneID:446823"
/translations="MYSGAGPLVPTTTPPLMPATPKPPPPDFGTSGRITKLOA
NPFEMDIPKIEIHYHIDIDIKPCPRVNRIVEHVVQHFPAQIFGDRKPFVDRKRL
YTAMPIDAKDQLEVLVLEGKDRIFKVAIKWACVSLQALHDLGRPLNVPFE
TVALDVWRHLSMRYTPVGRSFFTASEGANPLGGREWFQGVSRPSLWKMML
NIDSVATYKAQVIEFMCQVDFKSIIEQQKPLDSORVKFKIKGLKVKRTHCG
QMKRKYCNVTRRPASHOTFPLQSGGOTVECTVAQYKDRHKLVLRYPLPCLQVG
OEKHTYLPLEVCNIVAGORCIKILTDNOSTWIRATARSAPDQREISKLMSASFN
TFPFREFGLMWKDMIDTGVLPQPSILYGGRSKAIATPVQGVDMRNKQPHTEIG
IKWALACAPQCTCEHLKTEQLRKISRDAGMPIQQQPCFKYAGQADSVPEMF
RLHNTYGLQLVVLPKGTVPYAEVRVGTDLVGMATQCVQMKVQRTTPTLSNL
CLKINVLKGLGVNLLPQGRPPVFQGVPIFLAGADVTHPPAGDGKPSIAAVGSMDAH
PNRYCAVQVQHRBIIODLSAMVELLIQFYKSTRFKPRTIIFVRDGVSEGFQOV
LHHELLAIREACIKLEKDYQPGITFLVQKRBHTRLFCTDRNVERCKSNIPAGTTVD
TKITHPSDFPILCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYLQCHTYVRCR
SVSIPATYIAHLVAPRAYHLVDKEHDSBSGHTSGQNGRDRQALAKAVQVHQDTL
RTMYFA"

ALIGNMENT Scores:
Pred. No.: 3,56e-55 Length: 3298
Score: 567.00 Matches: 245
Percent Similarity: 40.1% Conservative: 156
Best Local Similarity: 24.5% Mismatches: 387
Query Match: 10.6% Indels: 213
DB: 11 Gaps: 39

US-10-645-746-3 (1-1020) x BC077863 (1-3298)

QY 11 GlyPheTyrArgHisSerLeuAspProGluMetLysThrLeuAlaArgProThrGlyLys 30
Db 325 GGCATATACCTTCAAGCCACCTCCAGCGGCAGATTTTGGTACCTCAGGGAGGACCATCAA 384

QY 31 CysAspGlyLys-PheTyrGluLysLysValLeuLeuLeuValAsnTrpPheLysPheSe 50
Db 385 GCTCCAGGCAACAGTTTGTGAATGACATT----- 415

QY 50 rSerLysIleTyrAspArgGluTyrTrpGluTyrGluValLysMetThrLysGluValLe 70
Db 416 -CCCAAAATT-----GAATATATCATGATGACATTGATATA----- 451

QY 70 uAsnArgLysProGlyLysProPheProLysLysThrGluIleProIleProAspArgAl 90
Db 452 -----AAACCAGAGAAA---TGCCCCAGACGAGTAACAGGGAATAGTGAACACAT 501

```

```

QY 90 aLysLeuPheTrpInHisLeuArgHisGluLysLysGlnThrAspPheIleLeuGluAs 110
Db 502 G-----GTCCAGCAGCTTTAAAGCTCAA-----ATCTTTGGGGA 534

QY 110 pTyr-----ValPheAspGluLysAspThrValTyrSerValCysArgLeuAsnThrVa 128
Db 535 TCGAAACACAGTGTGTTGATCGAAGAAAAAACCTTTATACTGCAATGCTCTGCCAAATGC 594

QY 128 lThrSerLysMetLeuValSerGluLysValValLysLysAspSerGluLysLysAspGl 148
Db 595 AAGAGATAACACAGGTG-----GAGTTAGAGTCACACTACCCGGCGAAGGAAGATCG 648

QY 148 uLysAspLeuGluLysLysIleLeuTyrThrMetIleLeuThrTyrArgLysLysPheHi 168
Db 649 T-----ATCTTCAAAGTTGCAATCAAGTGAATGATGTAAGCTTGCAGGACCTGCA 702

QY 168 sLeuAsnPheSerArgGluAsnProGluLysAspGluGluAlaAsnArgSerTyrLysPh 188
Db 703 CGAGCTCTGTCTGGTGGTCTTCCCAATGTCCTTTTGAG-----ACAGTTCAGGC 753

QY 188 eLeuLysAsnValMetThrGln-----LysValArgTyrAlaProPheValAsnGluGl 206
Db 754 TTTAGAGCTGTGATGATGAGACATTTACCTTCTATGAGGTATACACCTGTC----- 802

QY 206 uIleLysValGlnPheAlaLysAsnPheValTyrAspAsnAsnSerIleLeuArgValPr 226
Db 803 -----GGTCGTTCTGTTTTTACTGTCATC 825

QY 226 oGluSerPheHisAspProAsnArgPheGluGlnSerLeuGluValAlaProArgIleGl 246
Db 826 TGAAGCTCGCGCAAAATCCT-----CTTGGGGGTGGTAGAGA 861

QY 246 uAlaTrpPheGlyIleTyrIleGlyIleLysGluLeuPheAspGlyGluProValLeuAs 266
Db 862 AGTTTGGTTGGTTCATCATCAGTCTGTACGACCTTCCCTC---TGGAAATATGATGCTCAA 918

QY 266 nPheAlaIleValAspLysLeuPheTyrAsnAlaProLysMetSerLeuAspTyrLe 286
Db 919 TATCGATGTGTCTGCACAGCATTTTACAAGCA-----CAACCATGATTAATGAATTTAT 972

QY 286 uLeuLeuIleValAspProGlnSerCysAsnAspAspValArgLysAspLeuLysThrLy 306
Db 973 GTGTGAAGTCTGTGATTTTAAAGACATCGAAGACAA----- 1009

QY 306 sLeuMetAlaGlyLysMetThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLe 326
Db 1010 -----CAGAAACCTCTTACCGATTTCCCAAGGGTTAAGTTTACCAAGAAAT 1056

QY 326 uGluAsnLeuLysLeuLysCysAlaGluValTrpAspAsnGluMetSerArgLeuThrGl 346
Db 1057 TAAAGGTTTAAAGTCGAAATAAC-TCACTGTGACCAATGAA-----ACGGA 1103

QY 346 uArgHisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGl 366
Db 1104 A-----ATACCGAGTTTGCATGTAAACAGGCGACCGACCATCAACATCCCTC 1157

QY 366 yLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeu-PheLysIleTyrG 386
Db 1158 TTCAGCAGGAGTGCAGACAGCGGTGAGTGTACCTGCGACAGTACTTCAAG-----G 1211

QY 386 luGluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyA 406
Db 1212 ACAGACACAAACATAGTTCTGCGTACCTCATTTACCATGTTTACAGGTGGGCGAGGAGC 1271

QY 406 lAlysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrL 426
Db 1272 AGAAACACACAGTACTCTCCCTTGTGAGGTGTGCAATATATGTGGCAGCCAAAGATGCA 1331

QY 426 yAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---ArgLysP 445
Db 1332 AGAACTTACGAGC-----AATCAGACCTCCACTATGATACGGGCAACTGCTCGATCTG 1385

```

```
QY 445 roHieaspTyrLysGluAenThrLeuLysMetLeuLysGluLeuAaspPheSerSerGluG 465
Db 1386 CTCAGATCCGAGGAAGAGATCAGTAAATTAATCGAAGTGCAGTTTAACTACTGATC 1445
QY 465 luLeuAasnPheValGluArgPheGlyLeuLysSerLysLeuGlnMetIleGluCysProG 485
Db 1446 CC-----TTTGTGCGAGAGTTTGGTATTAATGGTGAAGATCAGATGCCGATGTCACAG 1499
QY 485 llyvalLeuLysGluProMetLeuValAasnSerValAasnGluGlnIleLysMetThrP 505
Db 1500 GCAGAGTCCTCAACCTCTTCAATCCTTTATGAGCGCAGGAGTAAAGCAATTGCACGCG 1559
QY 505 roVal-----lleargGlyPheGlnGluLysGlnLeuAasnVal---ValProGluLysG 522
Db 1560 CTGTTCCAGGTCGTGGGATATGAGAAACAACAGTTCACACCCGCGCATTTGAAATTAAGG 1619
QY 522 luLeuCysCysAlaValPheValAasnGluThrAlaGlyAasnProCysLeuGluGluA 542
Db 1620 TGTGGGCCATCGCTGCTTTGCCCCACAA-----CGACAGTGTACTGAAGTTC 1667
QY 542 enAaspValLysPheTyrThrGluLeuIle-Gly-----GlyCys 555
Db 1668 AC-----CTGAAGACATTTACAGAACAGTTCGAGGAAGATTTCTCGTGATGCCGGGATGC 1721
QY 556 LysPheAtgGlyleargIleGlyAlaAasnGluAasnAtgGlyAlaGlnSerIle----- 573
Db 1722 CCATCCAGGGGAGCGCATGTTCTGCAGATAT-GCCAGGGAGCAGATAGTGTGAGCCA 1780
QY 574 MetTyrAaspAlaThrLysAasnGluTyrAlaPheTyrLysAasnCysThrLeuAasnThrGly 593
Db 1781 ATGTTCCAGGCATCTGAAGATACATAT-----ACAGGG 1813
QY 594 lleGlyArgPheGluIleAlaAlaThrGluAlaLysAasnMetPheGluArgLeuProAasp 613
Db 1814 CTA----- 1816
QY 614 LysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAasnAlaTyrGly 633
Db 1817 -----CAACTGTTGTTGAATACTGCTCGGGAAACCCCTGCTATGGG 1861
QY 634 PheValLysHieTyrCysAaspHisThrIleGlyValAlaAasnGlnHisIleThrSerGlu 653
Db 1862 GAAGTGAACCGTGTGGGACACACTGTTCTGGGAATGGCTACTCAGTGTGTTCCAGATGAAA 1921
QY 654 ThrValThrLysAlaLeuAalaserLeuArgHisGluLysGlySerLysArgIlePheTyr 673
Db 1922 AATGTG-----CAGAGACCACACACCAAAACCTTGCT 1954
QY 674 GlnIleAlaLeuLysIleAasnAlaLysLeuGlyIleAasnGlnLeuAaspTrpSer 693
Db 1955 AACCTGTGCTTAAATAATAATGTAATAATTAGAGGGCTCAATAACATTTTG----- 2005
QY 694 GluIleAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMet 713
Db 2006 -----TTGCCCCAGGCGACACCTCTGCTGTTTCAACAGGCAGTCA---ATA 2047
QY 714 TyrValGlyIleAaspValThrHisProThrSerTyrSerGlyIleAaspTyrSerIleAla 733
Db 2048 TTTCTGGAGCTGATGTTTACCACCCCTGCAGGGGACGGTAAAGACCTTCGATGCT 2107
QY 734 AlaValValAlaSerIleAasnProGlyGlyThrIleTyrArgAasnMetIleValThrGln 753
Db 2108 GGGGTTGTAGTATGATGATGCACACCCCAACCGATATGTGCCATCTGGGGGTGCAG 2167
QY 754 GluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAaspIleLeu 773
Db 2168 CAG-----CACCAGCAAGAAATCATTCAGGATTTG 2197
QY 774 GluAlaLysPheValLysLeuLeuArgGluPheAlaGluAasnAasnAaspAasnArgAlaPro 793
Db 2198 TCTGCATGTTGGGAGCTGCTATCCAGTTCTCAACAGTCTACTGCTTCAAG---CCC 2254
QY 794 AlaHisIleValValTyrArgAaspGlyValSerAaspSerGluMetLeuArgValSerHis 813
```

```
Db 2255 ACNAGAATCATCTTTTACAGAGATGGTCTCTCAGGAAGGCGAGTCCACAGGTTCTCCAT 2314
QY 814 AspGluLeuAArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAaspGlyGlu 833
Db 2315 CATGAGCTTCGGCCATTTAGGGAAGCCCTGTATTAAACTG-----GAGAAG 2359
QY 834 AaspProGluProLysTyrThrPheIleValIleGlnLysAatqHisAasnThrArgLeuLeu 853
Db 2360 GATTACACGCGCCAGGAATTCATTTATTGTAGTGCAGAAAGACATCACACGCGCTATTTC 2419
QY 854 ArgAzmMetGluLysAaspLysProValValAasnLysAaspLeuThrProAlaGluThrAasp 873
Db 2420 ---TGTCAGACCGAAATGAGAGGGTAGGAAAGTGGAAATATACCACGT----- 2467
QY 874 ValAlaValAlaAlaValLysGlnTrpGluGluAaspMetLysGluSerLysGluThrGly 893
Db 2467 ----- 2467
QY 894 lleValAasnProSerSerGlyThrThrValAaspLysLeuIleValSerLysTyrLysPhe 913
Db 2468 -----GGTACCACCTGTAGACACAAAGATTACCACCCCTTCAGAGTTT 2509
QY 914 AaspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThr 933
Db 2510 GATTTTACCTGTGCAGTCAGCTGCTGTTATTCAGGGAACAAGCAGCGTCACACTATCAT 2569
QY 934 ValMetTyrAaspAaspLysGlyMetSerGlnAaspGluValTyrLysMetThrTyrGlyLeu 953
Db 2570 GTCTCTGTGGGATGACAATCGCTTTTCATCTGTATGATGTTGCAGATCTCTCATTCAGCTC 2629
QY 954 AlaPheLeuSerAlaAargCysArgLysProIleSerLeuProValProValHisTyrAla 973
Db 2630 TGCCACACTATGTTGTTGCACTCGTCTGTCTGTCTATCCCTGCTCCAGCATATATGCT 2689
QY 974 HisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db 2690 CACCTGTGGTTCATTCAGAGCAAG--TATCATTTTGGTGGATAAAGAGCAT 2737

RESULT 10
AX575501 3580 bp DNA linear PAT 07-JAN-2003
LOCUS Sequence 17 from Patent WO02072630.
DEFINITION AX575501
ACCESSION AX575501
VERSION AX575501.1 GI:27552093
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Thornton,M., Hafalia,A.J., Lu,D.A., Arvizu,C., Swarnakar,A., Lu,Y.,
Warren,B.A., Baughn,M.R., Tang,Y.T., Lee,E.A., Yao,M.G.,
Ramkumar,J., Khan,F.A., Gandhi,A.R., Ding,L., Yue,H., Gietzen,K.J.,
Walia,N.K., Thangavelu,K., Elliot,V.S. and Marquis,J.P.
Nucleic acid-associated proteins
Patent: WO 02072630-A 17-19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..3580
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1725129CB1"

ORIGIN
Alignment Scores:
Pred. No.: 5,26e-55 Length: 3580
Score: 566.00 Matches: 239
Percent Similarity: 40.0% Conservative: 159
Best Local Similarity: 24.0% Mismatches: 385
Query Match: 10.6% Indels: 214
```

DB: 2 Gaps: 38
US-10-645-746-3 (1-1020) x AX575501 (1-3580)
QY 19 ProGluMetLysTrpLeuAlaArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
DB 95 CCCATCCAAGGATATGCTTCAAGGCTCCACCTAGACCGGACTTTGGGACCTCCGGGAGA 154
QY 39 LysValLeuLeuValAsnTrpPheLysPheSerLysLysIleTyrAspArgGluTyr 58
DB 155 ACAATCAAAATTACAGGCCAAATTTCTCGAAATGGACATCCCCAAAATTGACATCATCAT 214
QY 59 TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhe 78
DB 215 TATGAAATGGATATCAAGCCAGAGAG-----TGC 244
QY 79 ProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArg 98
DB 245 CCGAGGAGAGTTAAACAGGGAATCTGGNAACACATG-----GTCCAG 286
QY 99 HisGluLysLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLys 116
DB 287 CACTTTAAACACAG-----ATCTTTGGGATCGGAAGCCCGTGTGTGAGCGCAGG 337
QY 117 AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
DB 338 AAGAAATCTATACAGCCATGCCCTTCCGATTGGGAGGCAAGGTG-----GAG 388
QY 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeu 156
DB 389 CTGGAGGTACCGTCCAGGAGAGGCAAGATCGC-----ATCTTCAAGGTGTCCATC 442
QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
DB 443 AAGTGGGTGCTCGTGTGAGCTTCCAGCGCTTACACGATGCACATTTTCAGGCGCGCTGCC 502
QY 177 GluLysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln--- 195
DB 503 AGCGTCCCTTTTGTAG-----ACGATCCAGGCCCTTGGACGTGGTCATGAGGCACCTG 553
QY 196 ---LysValArgTyrAlaProPheValAsnGluGluIleLysValGlnPheAlaLysAsn 214
DB 554 CCATCCATAGGTATACACCCCGTG----- 577
QY 215 PheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg 234
DB 578 -----GGCCGCTCTCTTCCAGCGGCTCCGAGGCTGCTTAACCT----- 619
QY 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrIleGly 254
DB 620 -----CTTGGCGGGCGGAGAGTGTGGTTGGCTTCCATCAGTCC 661
QY 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
DB 662 GTCCGGCTTCTCTC---TGGAAATGATGCTGAATATTGATGTCAGCAACAGCGTTT 718
QY 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrIleLeuLeuIleValAspProGlnSer 294
DB 719 TACAAGGCA-----CAGCCAGTAATCGAGTTTGTGTAAGTTTGGATTAAAAAGT 772
QY 295 CysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
DB 773 ATTGAAGAACAA-----CAAAAACCTCTG 796
QY 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys----- 332
DB 797 ACAGATCCCAAGGGTAAAGTTTACCAAGAAATTAAGCTTAAAGGTGGAGATAACG 856
QY 333 ---CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGluArg 347
DB 857 CACTGTGGGAGATGAGAGAGAGTACCGGCTCTGCAATGTGACCCCGCGCGCCAGT 916
QY 348 HisLeuThrPheLeuAspLeuCysGluGluAsnSerLeuValTyrLysValThrGlyLys 367

DB 917 CACCAAACTTC---CCGCTGCAGCAGGAGCGGGCAGACGGTGGAGTGCACG----- 967
QY 368 SerAspArgGlyArgAsnAlaLysLysTyrAspThrThrLeuPheLysIleTyrGluGlu 387
DB 968 -----GTGGCCCATGAT-----TTCAAG-----GACAGG 991
QY 388 AsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAlaLys 407
DB 992 CACAAGTTGGTTCTCGCTACCCCACTCCCATGTTTACAAGTCGGACAGGACAGAAA 1051
QY 408 GluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAsn 427
DB 1052 CACACCTACCTTCCCTGGAGGTCTGTAACATTTGGCAGCAGCAAAAGATGTATTAAAAA 1111
QY 428 ArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr-----ArgLysProHis 446
DB 1112 TTAACGGAC-----AATCAGACCTCAACATGATCAGAGCGACTGTAGTTCGGCGCC 1165
QY 447 AspTyrLysGluAsnThrLeuLysMetLysGluLeuAspPheSerSerGluGluLeu 466
DB 1166 GATCGGCAAGAAGAGATTAGCAATTGATCGAAGTCAAGTTTCAACACAGATCCA--- 1222
QY 467 AsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLys 486
DB 1223 ---TACGTCCGTGAATTTGGAATCATGTCANAGATGAGATGACAGACGTGACTGGCGG 1279
QY 487 ValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrProVal 506
DB 1280 GTCTGCGCGCGCTCCATCTCTACGGGGGCGAGGAATTAAGATATTTCGACCCCTGTG 1339
QY 507 -----IleArgGlyPheGlnGluLysGlnLeuAsnVal---ValProGluLysGluLeu 523
DB 1340 CAGGGCTCTGGACATCGGAAACAGCAGTTCCACACGGGCGATCGAGATCAAGGTGG 1399
QY 524 CysCysAlaValPheValValAsnGluThrAlaGlyAsnProCysLeuGluGluAsnAsp 543
DB 1400 GCATTCGCTGCTTGGCCCCCAG-----CGCCAGTGCACGGAAGTCCAT--- 1444
QY 544 ValValLysPheTyrThrGluLeuIle-----GlyGly-CysLysPhe 557
DB 1445 ---CTGAGTCTTTCAGAGAGAGCTCAGAAAGATCTCGAGAGAGCGTGGATCCCATC 1501
QY 557 eArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSerIle-----MetTy 575
DB 1502 CAGGGCCAGCGGTCTTCTGCAAAATACGCG-CAGGGGGCGACAGCGTGGAGCCCATGTT 1560
QY 575 rAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyIleG 595
DB 1561 CCGGCACCTTGAAGAACACGATATCG----- 1585
QY 595 YArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLysG 615
DB 1585 ----- 1585
QY 615 uGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPheVa 635
DB 1586 ---GGCCTGCAGCTGTGTGTGTCTCTGCGCGGCAAGACCGCCGTCACGCCGAGGT 1641
QY 635 lLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluThrVa 655
DB 1642 CAAGCGCTGGGAGACACGCTGCTGGGATGGCCACGCTGCGTCAGATGAAGAACGT 1701
QY 655 lThrLysAlaLeuAsnSerLeuArgHisGluLysGlySerLysArgIlePheTyrGlnI 675
DB 1702 G-----CAGAGACCGCCACACACCTGTGTCACCT 1734
QY 675 eAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpSerGluI 695
DB 1735 CTGCTTGAAGATCAACGTCAAGCTGGGAGCGGTGAACACATCTCTG----- 1780
QY 695 eAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTyrVa 715

```
Db 1781 -----CTGCCCCAGGGCAGCGCGCGGTGTTCACGACGCGGTC---ATCTTTCT 1827
QY 715 lGlylleAspValThrHisProThrSerTyrSerGlylleAspTyrSerIleAlaAlaVa 735
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1828 GGGAGCAGACGTCACCTCCACCCCGCGGGATGGGAAGAAGCCCTCCATTGCGCGCGT 1887
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 735 lValAlaSerIleAenProGlyGlyThrIleTyrArgAenMetIleValThrGlnGluGl 755
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1888 GTGGGCAGCATGGAGCCGCCACCCCATCGCTACTCGCCACCGTGGCGGTGCAGCAG-- 1945
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 uCyBArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGluAl 775
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1946 -----CACCGGCAGGAGATCATACAGACCTGGCGCG 1977
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 alySpheValValysLeuLeuArgGluPheAlaGluAenAenAspAenArgAlaProAlaHi 795
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1978 CATGTGTCGCGAGCTCTCATCCAGTCTTACAAGTCCACGCGCTTCAAG---CCCACCG 2034
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 795 sIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGl 815
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2035 CATCATCTTACCGCGAGCGGTCTCTGAAGCCAGTTCACGACGTTCTCACACCA 2094
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 uLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPr 835
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2095 GTTGTGGCCATCCGTGAGCGCTGTATCAAGCTA-----GAAAAAGACTA 2139
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 oGluProLysTyrThrPheIleValleGlnLysArgHisAenThrArgLeuArgAr 855
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2140 CCAGCCCGGAGTACCTTCATCGTGTGCAGAGAGGCACACCCCGGCTTCTC 2196
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 gMetGluLysAspLysProValValAenLysAspLeuThrProAlaGluThrAspValAl 875
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2197 CACTGACAAAGACGAGCGGGTGGGNAAGTGGAAACATTCACGCA----- 2242
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 aValAlaValLysGlnTrpGluLysGlnMetLysGluSerLysGluThrGlyIleVa 895
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2242 ----- 2242
QY 895 lAenProSerSerGlyThrValAspLysIleuLeuValSerLysTyrLysPheAspPh 915
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2243 -----GGCACACACTGTGGACACGAAATCACCCACCCACCGAGTTCGACTT 2289
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 915 ePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMe 935
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2290 CTACCTGTGTATCAGCTGGATCCAGGGGACAGAGCGCTTCGCACTATCAGCTCT 2349
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 tTyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPh 955
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2350 CTGGGACGACAATCGTTTCTCTCTGATGAGCTGCAGATCCTAAACCTACGCTGTGTC 2409
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 955 eleuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHisle 975
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2410 CACCTACGTGCGCTGCACACGTCGCTGTCATCCAGCGCCAGCATACGCTCACCT 2469
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
QY 975 uSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2470 GTGGCCTTCGGGCGCAGG---TACCACCTGTGTGATAGGAACAT 2512
   :||| ||||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
LOCUS BC075263
DEFINITION Xenopus tropicalis MGC88879 protein, mRNA (CDNA clone MGC:88879
IMAGE:6988277), complete cds.
ACCESSION BC075263
VERSION BC075263.1 GI:49522951
KEYWORDS MGC.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 3687)
AUTHORS Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
```

Collins FS, Wagner L, Shennan CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsien F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Ustin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McSwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettman M, Madan A, Rodriguez S, Sanchez A, Whitting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA. Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3687)

Klein,S. and Gerhard,D.S.

Direct Submission

Submitted (28-JUN-2004) National Institutes of Health, Xenopus Gene

Collection (XGC), National Institute of Child Health and Human

Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD

20892-7510, USA

NIH-MGC Project

Contact: XGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Rob Granger, University of Virginia

cDNA Library Preparation: Open Biosystems

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,

Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth

Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,

Kim MacDonald, Amara Maason, Mike R. Mayo, Josh Moran, Ryan Morin,

Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prähbu,

Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska,

Duane Smalls, Jeff Stott, Miranda Tsai, George Yang, Jacque

Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 171 Row: a Column: 14

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 52345659.

Location/Qualifiers

1. 3687

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="MGC:88879 IMAGE:6988277"

/tissue_type="whole body, male, 10 months old, strain F6,

normal"

/clone_lib="NICHD_XGC_SwblN"

/lab_host="DH10B"

/note="Vector: pExpress1"

1. 3687

/gene="MGC88879"

/db_xref="GeneID:448205"

239..2854

/gene="MGC88879"

/codon_start=1

/product="MGC88879 protein"

/protein_id="AAH75263.1"

/db_xref="GI:49522952"

/db_xref="GeneID:448205"

/translation="MYAGAGFVLVPPTTPPLMPAYTFKPPRPDQCTSGRTIKLOA

NFFENDIPKIEIHYEIDIKPEKPRVRNREIHHVQHFKAQIFGDRAPQVDRGNL

YTAMPLPTARDKQVEVLTPFEGEKDRIKFKVAIKWMACVSLQALHDAUSGRLPSPVFE

TIOALDVMRHLPSMRYTPVGRSPFTASGECANPLGGREWFEGFHSVRSPLWKWML
 NIDVSAFYKAQPIVIFEMCEVLDPKSIEBQQRKLQDSQVRFKTEIKGLKVELTHCG
 QMKKRYCNVTRPASHQTPLOQESGQVECTVAQYFKDRHKLVLURYPHLPCLQVG
 QEQHTLPLEVCNIVAGQRCIKGLTDNQTSTWIRATARSAPDRQEEISKLMSRSFN
 TDPFREFGIMVKDDMTDVTGRVLQPPSILYGGVWEPNAPLNKALATVPQGVWDMR
 NKQPHTEIEIKWAIACFAPORQCTEVLKTFTEBLRKISRDAGMPLOQPCFKYAO
 GADSVEMFRLHNTYTGLOLVVILPGKTPVTAERKRGDTVLGMATOCVOMKNVOR
 TTPQTLNCLKLNKLVGGVNNLLPQGRPPVFOQPIVPLGADVTHPPAGGKKPSIA
 AVUGSDMHPNRICATVRVQHQREIILQDSNMRRELLIQFYKSTRFKPRLIIFYRDG
 VSGQFOVLHHELLAREACIKLEKDYQFETIVVQKRHHHTLFTDRNRVKGSG
 NIPAGTVDTDTKTHPSFDFGLSHAGIOGTSRPSHYHVLWDDNRFSSDLQILTYOL
 CHTYVCTRSVSIIPAPAYYAHVAFRARYHLVAFRHYLHDKHDSARGSHTSGSNGRDQOALAK
 AVQVHQDTLRTMYFA"

ORIGIN

Alignment Scores:

Pred. No.: 5,49e-55 Length: 3687
 Score: 566.00 Matches: 247
 Percent Similarity: 40.2% Conservative: 157
 Best Local Similarity: 24.6% Mismatches: 380
 Query Match: 10.6% Indels: 224
 DB: 11 Gaps: 40

US-10-645-746-3 (1-1020) x BC075263 (1-3687)

QY 19 ProGluMetLysTrpLeuAlaAArgProThrGlyLysCysAspGlyLysPheTyrGluLys 38
 DB 293 CCAATGCCGCGGTATACCTTCAGACCTCCGCCACGCCGACGATTTTGGTACCTCAGGGAGG 352
 QY 39 LysValLeuLeuValAsnTrpPheLysPheSerLysIleTyrAspArgGluTyr 58
 DB 353 ACCATCAAGCTCCAGCCCACTTTTGAATGGAC-----ATTCCAAAATTGAATA 406
 QY 59 TyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysProPhe 78
 DB 407 TATCACTATGAGATTGATATA-----AAACCAGAGAAA---TGC 442
 QY 79 ProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeuArg 98
 DB 443 CCAGACGAGTAACAGGGAATAGTTGAACACATG-----GTCCAG 484
 QY 99 HisGluLysLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGluLys 116
 DB 485 CACTTTAAAGCTCAA-----ATCTCGGGAGTCGAAACACGATGTTTGAAGAGA 535
 QY 117 AspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSerGlu 136
 DB 536 AAAAACTTTACACTGCAATGCCTCTCCGATTGCAAGAGATAACAGGTG-----GAG 589
 QY 137 LysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIleLeu 156
 DB 590 TTAGAAGTCACACTACTGGGGAAGGGAAGATGCG-----ATCTTCAAAGTTGCAATC 643
 QY 157 TyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsnPro 176
 DB 644 AAGTGATGCGGTGTAGCTTGACGGCCCTACATGATGCACTGCTGTGCTCTTCCC 703
 QY 177 GluLysAspGluGluAlaAsnArgSerTyrLysPheLysAsnValMetThrGln--- 195
 DB 704 AGTGTCCTTTTGAG-----ACAAATTCAGGCTTTAGATGTTGTGATGAGGCATTTA 754
 QY 196 ---LysValArgTyrAlaProPheValAsnGluLysValGlnPheAlaLysAsn 214
 DB 755 CCTTCTATGAGGTATACACAGTT----- 778
 QY 215 PheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsnArg 234
 DB 779 -----GGTCGTCTTTTCTTACTGCACTCTGAAGCTGTGCAATCTC-----820
 QY 235 PheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrGly 254
 DB 821 -----CTCGGGGTGGCAGAGAGTTTGGTTTGGATTTCATCAGTCC 862

QY 255 IleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeuPhe 274
 DB 863 GTACGCCGCTGCTC---TGGAAAATGATGCTCAATATCGATGCTCTGTACACGACATT 919
 QY 275 TyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuIleValAspProGlnSer 294
 DB 920 TACAAAGCA-----CAACCAAGTANTTGAAGTTTATGTGTGAAGTTCTGATTTTAAAGC 973
 QY 295 CysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThrIle 314
 DB 974 ATCAGAGGAGCA-----CAGAAAACCTCTG 997
 QY 315 ArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysCysAla 334
 DB 998 ACTGATTTCCAAAGGGTTAAATTTTACCAAGAAATTTAAAGGTTTAAAGTGGAAATAC- 1056
 QY 335 GluValTrpAspAsnGluMetSerArgLeuThrGluArgHisLeuThrPheLeuAspLeu 354
 DB 1057 TCACGTGGACAATGAA-----ACGGAATATACAGAGT 1089
 QY 355 CysGluGluAsnSerLeuValTyrLysValThrGlyLysSerAsp----- 369
 DB 1090 TTGCAACGCTAAC-----AAGCGCACCGACGAGTCATCAACATTCCTCTCTCCA 1137
 QY 370 -----ArgGlyArgAsnAlaLysLysTyrAspThrThrLeu-PheLysIleTyrGlu 387
 DB 1138 GCAGGAGAGTGGACAGACGCTAGATGACCGCTGCACAGTACTTTAAG-----GACAG 1191
 QY 387 uAsnLysLysPheIleGluPheProHisLeuValLysValLysSerGlyAlaLys 407
 DB 1192 ACACAACTAGTCTCGGTTACCTCATTTTACCATGTTTACAGTGGCGGAGGAGCAGAA 1251
 QY 407 sGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyrLysAs 427
 DB 1252 ACACACATACCTCCACTTGGAGTATGCAATATAGTGGCTGGTCAAGAGTCATAAGAA 1311
 QY 427 nArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---ArgLysProHis 446
 DB 1312 ACTTACGGAC-----AATCAGACCTCCATATGATACGGGCACTGCTCGATGTGCTCC 1365
 QY 446 sAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGluLe 466
 DB 1366 AGATCCCAAGAAAGACATCAGTAATTAATGCGAAGTGCAGGTTTANTACTGATCCC-- 1423
 QY 466 uAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysProGlyLys 486
 DB 1424 ---TTGTGCGAGAGTTTGGTATAATGTTGTAAGATGACATGACCGATGTCACAGGCG 1479
 QY 486 sValLeuLysGluProMetLeuVal-----AenSerVa 497
 DB 1480 AGTCTTCAACCTCTTCAATCCTTTATGAGGAGGAGGTATGGGAAGAGCAAAATGCTCC 1539
 QY 497 LaenGluGlnIleLysMetThrProVal-----IleArgGlyPheGlnGluLysGlnLe 515
 DB 1540 CTTGAACAAAGCAATTCGCCACCTCTGTCAGGGTGTCTGGGACATGAGAACAAACAGTT 1599
 QY 515 uAsnVal---ValProGluLysGluLeuCysCysAlaValPheValValAsnGluThrAl 534
 DB 1600 CCACACCGGCATTTGAAATAAGGTGTGGGCCATCGCTGCTTGGCCCAAC----- 1651
 QY 534 agLysAsnProCysLeuGluGluAsnAspValValLysPheTyrThrGluLeuIle-Gly- 553
 DB 1652 ----CGGCAGTGTACTGAAGATTTCAC-----CTCAAGACATTTTACAGACAGTTGAGAA 1701
 QY 554 -----GlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsn 568
 DB 1702 GATCTCAGCTGATGCGGGATGCTTATCCAGGACAGCGCTGTTCTTGCATAATAT-GCCC 1760
 QY 568 rgGlyValGlnSerIle-----MetTyrAspAlaThrLysAsnGluTyrAlaPheTyrL 586
 DB 1761 AGGAGCAGATAGCGGTGAGCCAAATGTTTCAGGCATCTGAAGATACTTAT----- 1810
 QY 586 yeAsnCysThrLeuAsnThrGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLys 606

Db	1811	-----ACAGGCTA-----	1819
Qy	606	snMetPheGluArgLeuProAspLysGluGlnLysValLeuMetPheIleIleIleSerL	626
Db	1820	-----CAACTAGTCGTTGTAATCTTC	1841
Qy	626	ysArgGlnLeuAenAlaTyrGlyPheValLysHisTyrCysAspHisThrIleGlyValA	646
Db	1842	CTGGGAAACCCCTGCTCTATCGGAAGTGAACGTTGGGGGACACTGTTCTCGGAATGG	1901
Qy	646	laAsnGlnHisIleThrSerGluThrValThrLysAlaLeuAlaSerLeuArgHisGluL	666
Db	1902	CTACTCAGTGTGTTGAGTGAATAATGTG-----	1934
Qy	666	ysGlySerLysArgIlePheTyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyGlyI	686
Db	1935	GGACCACACCAACCTTGTCTTAACCTGTGCTTAAATAATCAATGTAAATTTAGGAGCG	1994
Qy	686	leAsnGlnLeuLeuAspTyrPsrGluIleAlaGluLeuSerProGluGluLysGluArgA	706
Db	1995	TCAATAACATCTTG-----TTGCCCCAGGCGACACCTCCTG	2030
Qy	706	rgLysThrMetProLeuThrMetTyrValGlyIleAspValThrHisProThrSerTyrS	726
Db	2031	TGTTTCAACAGCTGTC-----ATATTCTTGGAGCTGATGTTACCCACCACAGCAGGG	2087
Qy	726	erGlyIleAspTyrSerIleAlaValAlaSerIleAsnProGlyGlyThrIleT	746
Db	2088	ACGGCAAAAGCCTTCAATTGCTGGGTTGTAGTAGTATGCACACCCCAACCGAT	2147
Qy	746	yrArgAsnMetIleValThrGlnGluGluCysArgProGlyGluArgAlaValAlaHisG	766
Db	2148	ACTGTGCCACTGTGAGGGTACAGCAG-----CACCA	2177
Qy	766	lyArgGluArgThrAspIleLeuGluAlaLysPheValLysLeuLeuArgGluPheAlaG	786
Db	2178	GGCAGAAATCATCCNGGATTTGTCTGCCATGTCGAGAGCTGCTTATCCAGTTCTACA	2237
Qy	786	luAsnAsnAspAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAsps	806
Db	2238	AGTCAACACGCTTCAAG-----CCACAGAATCATCTTTTACAGACGCGTGTCTCGAAG	2294
Qy	806	erGluMetLeuArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnP	826
Db	2295	GACAGTTTCAACAGGTGCTCCATCATGAGCTTCTGCCATTAGGGAAGCCTGATTAAAC	2354
Qy	826	heMetSerGluArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnL	846
Db	2355	TG-----GAGAAGGATTACACCGCGGAATTACATTTATTGTAGTGCAAA	2399
Qy	846	ysArgHisAsnThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysA	866
Db	2400	AAAGGCATCACACACGCTATTTC-----TGTAACAGCCGAATGAGCGAGTAGGGAAGTG	2456
Qy	866	spLeuThrProAlaGluThrAspValAlaValAlaValLysGlnTyrPgluGluAspM	886
Db	2457	GAATATATACAGCC-----	2470
Qy	886	etLysGluSerLysGluThrGlyIleValAsnProSerSerGlyThrValAspLysL	906
Db	2471	-----GGTACCACCGTAGACACAA	2489
Qy	906	euIleValSerLysTyrLysPheAspPheLeuAlaSerHisHisGlyValLeuGlyT	926
Db	2490	AGATACCCACCCATCAGAGTTTGTATTTTACTGTGCAGTCACGCTGTTATTCAGGAA	2549
Qy	926	hrSerArgProGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAspGluV	946
Db	2550	CAAGCAGGCGCTGCACACTATCATGCTCTGGGATGACAAATCGCTTTTTCATCGGATGAC	2609
Qy	946	alTyrLysMetThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerL	966

Db 2610 TGCAGATCCTCACTTACCAGCTCTGCCACACTTATGTGCGTTGCACTCGCTCTGTATCTA 2669

Qy 966 euProValProValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgT 986

Db 2670 TCCCTGCTCCAGCATACTATGCTACCTGCTGGTGCATTCAGACGAAGG---TATCATTTGG 2726

Qy 986 hrTyrLysGluHis 990

Db 2727 TGGATAAAGAGCAT 2740

RESULT 12

LOCUS AB081472 2657 bp mRNA linear ROD 24-JAN-2003

DEFINITION Mus musculus mRNA for Piwi/Argonaute family protein meif2C2, complete cds.

ACCESSION AB081472

VERSION AB081472.1 GI:22830888

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 Doi, N., Zenko, S., Ueda, R., Ohki-Hamazaki, H., Ui-Tei, K. and Saigo, K. Short-Interfering-RNA-Mediated Gene Silencing in Mammalian Cells Requires Dicer and eIF2C Translation Initiation Factors Curr. Biol. 13 (1), 41-46 (2003)

REFERENCE 2 (bases 1 to 2657) Doi, N. Direct Submission

TITLE Submitted (12-MAR-2002) Noboru Doi, University of Tokyo, Graduate School of Science, Department of Biophysics and Biochemistry, Yayoi 2-11-16, Bunkyo-ku, Tokyo 113-0032, Japan (E-mail: s816337@mail.ecc.u-tokyo.ac.jp, Tel:81-3-5841-4404, Fax:81-3-5841-4400)

FEATURES Location/Qualifiers

1..2657 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /cell_lines="NIH3T3" /cell_type="adherent" 1..2657 /gene="meif2C2" 28..2610 /gene="meif2C2" /function="siRNA-mediated RNAi?" /experiment="experimental evidence, no additional details recorded" /note="PRP domain, PAZ domain, PIWI domain" /codon_start=1 /product="Piwi/Argonaute family protein meif2C2" /protein_id="BAC15767.1" /db_xref="GI:22830889" /translation="MYSAGPVLASPAPTSIPGVAKPPPPDPGTTGRTIKLOAN FFEIDIPKIDIIHYEVDIKPEKPRRVREIHEVMVQHFKTOIFGDRKPFVFGKMLY TAMLPIDGRKVELEVDLPGEKDRILKVSINWVSQALHDALSLPSVPFETI QALDVVHMLPSMRYPVGRSFFFTASGNSPLPGGREVFGFHSVRPSLWKMMLNI DVSATVAFYKAQVPIEFCEVLDPKSIEEQKPLTDSQRVKFTKEIKGLKVEITHCCQM KRKYVNCVTRRPAASHOTFPLQOESGQTVECTVAQVFKDRHKLVLRYPHLPLCLQVGO QKHYLPLEVNCIVAGRCIKLUTDQTSMTIATARSAPORQEISIKLMRSASFNPD PYRFRGIMVNDMDTGTVRVLQPPSILYGRNKALATPVQSWDMRNKQFHTGLEIK VWTACFAPQRCQTEVHLKSFTQKRISRDAGMPIQGPCFKYAGGADSVEPMFRH LKNTYAGLQVVVILPGKTPVYAEVRGVDVLGMATQCVQMKNPSIAAVVGSMDAHPN KINKLGGVNNIILPQGRPPVFPQVPIFLGADVTPPAGDGKPSIAAVVGSMDAHPN RYCATTRVQOHOEIIODLAAMVRELLIOFYKSTRFKPTRIIFYRDKGVSEGOQOVLH HELLAIRACIKLEKDYOGTFLVQKHHTLFTCDKNERVGSGLNIPAGTVDTK ITHPTFDFYLCSHAGIOGTRPSHYHVDNDRFSSDELOLITLQOLCHTYVCTRFSV SIPAPYAHILVAFRARYHLVDKEHDSAEGSHTSGSNGRDRHQALAKVQVQDILRT MYFA"

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:	4.45e-55	Length:	2657
Score:	565.00	Matches:	231
Percent Similarity:	39.2%	Conservative:	159
Best Local Similarity:	33.2%	Mismatches:	355
Query Match:	10.6%	Indels:	250
DB:	6	Gaps:	40

US-10-645-746-3 (1-1020) x AB081472 (1-2657)

38 LysLysValLeuLeuLeuValAsnTrpPheLysPheSerSerLysLysLeuTyrAspArgGlu 57
136 AGAACAAATCAAAATCAGGCCCAATTTCTTGAATGGACATCCCAAAATTTGACATCTAT 195
58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro 77
196 CACTATGAATTGGACATCAACCTCAGAAACGC----- 228
78 PheProLysLysThrGluLeuProLeuProAspArgAlaLysLeuPheTrpGlnHisLeu 97
229 ---CCTCGAGAGTAGACCGGAAATTTGGAGCACATG-----GTC 267
98 ArgHisGluLysGlnThrAspPheLeuLeuGluAspTyr-----ValPheAspGlu 115
268 CAGCACTTTAAACTCAG-----ATCTTCGGGACCGGAAGCCAGTGTTTGATGGA 318
116 LysAspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSer 135
319 AGGAAGAATCTGTACACAGCAATGCCCTCCGATCGCAGGAGCAAGGTG----- 369
136 GluLysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysLeu 155
370 GAGCTGGAGGTACACCTCCCGGAGAGGCAAAAGAT---CGCATCTTTAAAGGTATCCATC 426
156 LeuTyrThrMetIleLeuThrTyrArgLysLysPheHisLeuAsnPheSerArgGluAsn 175
427 AAGTGGGTGTCGTGCTGAGCTG---CAGCGCTTACACGATGCACCTTCGGGGCGGCTG 483
176 ProGluLysAspGluGluAlaAsnArgSerTyrLysPheLeuLysAsnValMetThrGln 195
484 CCCAGCGTCCCTTCGAG-----ACGATCCAGCCCTCGACGTTGTTCATGAGGCAC 534
196 -----LysValArgTyrAlaProPheValAsnGluLeuLysValGlnPheAlaLys 213
535 TTACATCCATGAGGTACACCCCTGTT----- 561
214 AsnPheValTyrAspAsnAsnSerIleLeuArgValProGluSerPheHisAspProAsn 233
562 -----GGCCGTTCTTCTTCACTGCATCTCAAGGCTGTTCACACCTT--- 603
234 ArgPheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrLe 253
604 -----CTGGTGGGGGCAGAGTGTGTGTTCCTTCATCAG 642
254 GlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeu 273
643 TCCGTCGCCACCTTCTCTT---TGAAAAATGATGCTGAATATTGATGTATCGGCAACAGCG 699
274 PheTyrAsnAlaProLysMetSerLeuLeuAspTyrLeuLeuLeuValAspProGln 293
700 TTTTACAAGGCA-----CAGCCAGTGTATCGAGTTGTTGTGAAGTTTGGATTTTAA 753
294 SerCysAsnAspAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThr 313
754 AGTATTGAAGAACAA-----CAAAACCT 777
314 IleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLysLeuLys--- 332
778 CTGACAGATTCCCAAGGGTAAAGTTTACCAAGAAATCAAAGGTCTAAAGGTGGAGATA 837
333 -----CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGlu 346

838 ACGCACTGTGGTTCAGATGAAGAGGAGTACCGTGTCTGCAATGTGACCCGGCGGCTGCC 897

347 ArgHisLeuThrPheLeuAspLeuCysGluGluAenSerLeuValTyLysValThrGly 366

898 AGTCAACCAACGTTCC---CCACTGCAGCAGGAGAGTGGCGACAGCAGTGAGTGTACA--- 951

367 LysSerAspArgGlyArgAsnAlaLysLysTyThrAspThrThrLeuPheLysIleTyrglu 386

952 -----GTGGCCCACTAG-----TTCAAG-----GAC 972

387 GluAenLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAla 406

973 AGGCCAAGCTGGTTCCTCGCTACCCCACTCCCGCTTTTACAAAGTCGGACAGAGCAG 1032

407 LysGluTyAlaValProMetGluHisLeuGluValHisGluLysProGluArgTyrgly 426

1033 AAACACACTCTCTTTTGAGAGTCTGTAAACATAGTTGCTGGACAGAGATGTATATAA 1092

427 AsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---ArgLysPro 445

1093 AAATTAACAGAC-----GATCAGACCTCAACCATGATCATGACAGCAACTGCCAGGTCAGCA 1146

446 HisAspTyrgLysGluAenThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGlu 465

1147 CCTGATGCCCAAGAGAGATCAGACAACTGATGCGAAGTGCAGATTTCAATACAGACCCA 1206

466 LeuAenPheValGluArgPheGlyLeuCysSerLysLeuGluMetIleGluCysProGly 485

1207 -----TATGTTCTGTGAATTTGGAAATCATGTTGAAAGATGAGATGACAGATGTGACTGGG 1266

486 LysValLeuLysGluProMetLeuValAenSerValAenGluGlnIleLysMetThrPro 505

1261 CGGGTCTCGACGCCCTCCATCTCTACGGGGCGGAGCAACAAGCAATGGCCACCCT 1320

506 Val-----IleArgGlyPheGlnGluLysGlnLeuAenVal----- 517

1321 GTCCAGGGTGTGTGGGCATCGGAAACAAGCAGTTCCACAGGCGATCGAGATCAAGGTG 1380

518 -----ValProGluLysGluLeuCysValAalValPheValValAen 531

1381 TGGGCCATCGCTCTCGCTCCCGCCAGGCCAG---TGTACAGAAGTCCATCTCAAGTCC 1437

532 GluThr-----AlaGlyAsnPro 537

1438 TTCACAGAGCAACTGAGGAAGATCTCGAGAGATGCTGGGATGCCGATCCAGGGTCAGCCC 1497

538 CysLeuGluGluAenAspValValLysPheTyThrGluIleLysGlyCysLysPhe 557

1498 TGCTTC-----TGTAAATAC 1512

558 ArgGlyIleArgIleGlyAlaAenGluAenArgGlyAlaGlnSerIle-----MetTyr 575

1513 GCA-----CAGGGTCACAGACTGTGGAGGCCCATGTTCC 1545

576 AspAlaThrLysAenGluTyAlaPheTyrgLysAenCysThrLeuAenThrGlyIleGly 595

1546 CGACACCTGAAGAACACATACGCT----- 1569

596 ArgPheGluIleAlaAlaThrGluAlaLysAenMetPheGluArgLeuProAspLysGlu 615

1569 ----- 1569

616 GlnLysValLeuMetPheIleIleIleSerLysArgGlnLeuAenAlaTyrgLysPheVal 635

1570 ---GGCTCCAGCTGTGGTGGTGCATCTCCCTGGCAAACTCCTGTGTATCGGAAGTC 1626

636 LysHisTyrgLysAspHisThrIleGlyValAlaAenGlnHisIleThrSerGluThrVal 655

1627 AAGCGTGTGGGAGACACAGTCTCGGGATGCCCCACACAGTGCCTCCAGATGAAGAAGCTG 1686

656 ThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePheTyrgIle 675

1687 -----CAGAGACACAGCCACACAGACCCCTATCCATCTC 1719


```
QY 676 AlaLeuValLeuAsnAlaLysLeuGlyGlyLeuAsnGlnGluLeuAspTrpSerGluLeu 695
Db 1720 TGCTTAAAGATCAATGTCAAATGGGAGCGTCAACACATCCTG----- 1764
QY 696 AlaGluLeuSerProGluGluLysGluAsgAsgLysThrMetProLeuThrMetTyrVal 715
Db 1765 -----CTGCCACAGGCGAGCGCTCCAGTGTTCACCAACCTGTC---ATCTCTCG 1812
QY 716 GlyLeuAspValThrHisProThrSerTyrSerGlyLeuAspTyrSerIleAlaVal 735
Db 1813 CGAGCGCATGTCAACCACCCAGCAGTGGGAGTGGGAAGAACCTTCTATTGCTGCGGT 1872
QY 736 ValAlaSerIleAsnProGlyThrIleTyrArgAsnMetIleValThrGlnGlu 755
Db 1873 GTGGCAGCATGGACGACACACCCCAACCGCTACTGTGCCACCGTGGCGGTGACGAGC--- 1929
QY 756 CysAsgProGlyGluArgAlaValAlaHisGlyAsgGluArgThrAspIleLeuGluAla 775
Db 1930 -----CACCGACAGGAGATCATCCAGGACCTGGCTGCC 1962
QY 776 LysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProLahis 795
Db 1963 ATGTTACGAGAGTTGCTCATTCAGTTCTCAAGTCCACCGCTTCAAG---CCACCCGC 2019
QY 796 IleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAspGlu 815
Db 2020 ATCATCTTACCGCATGCGCTGCTCCGAGGCGCATGTTCCACGAGGTTCTCCACCATGAG 2079
QY 816 LeuAsgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAspPro 835
Db 2080 CTCCTGGCCATCAGAGCGCTGCATCAGCTG-----CAGAGGACTAT 2124
QY 836 GluProLysTyrThrPheIleValIleGlnLysAsgHisAsnThrArgLeuLeuArgArg 855
Db 2125 CAGCCAGGAATCAGTTTCTGTTGTCGAGAGCGGCACACACACGCTCTTC---TGC 2181
QY 856 MetGluLysAspLysProValValAsnLysAspLeuThrProAlaGluThrAspValAla 875
Db 2182 ACAGACAAAATGAGCGGTGTGGAGAGTGGGAACATTCCTCCGCA----- 2226
QY 876 ValAlaAlaValLysGlnTrpGluGluAspMetLysGluSerLysGluThrGlyIleVal 895
Db 2226 ----- 2262
QY 896 AsnProSerSerGlyThrThrValAspLysLeuIleValSerLysTyrLysPheAspPhe 915
Db 2227 -----GGCACAAACCGTGGACACGAAGATCACCCACCCCACTGAGTTGACTTC 2274
QY 916 PheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisTyrThrValMet 935
Db 2275 TACCTGTGCAGTCACCGCGCATCCAGGGACAGGCGGCTTCCCACTACCACTGCTT 2334
QY 936 TyrAspAspLysGlyMetSerGlnAspGluValTyrLysMetThrTyrGlyLeuAlaPhe 955
Db 2335 TGGGATGACAATCGTTTTTCTCTGATGAGCTGCAGATCTTGACCTACCACTGCTGCAC 2394
QY 956 LeuSerAlaArgCysArgLysProIleSerLeuProValProValHisTyrAlaHisLeu 975
Db 2395 ACCTAGTGTGTCACAGCTCTGTGTCAATACCGCGCCAGCTTACTATGCTACCTG 2454
QY 976 SerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGluHis 990
Db 2455 GTGGCCTCCGAGCCAGG---TACCACCTGGTGGATAAGAATC 2496
RESULT 13
AF005355 3599 bp mRNA linear MAM 25-JUN-1998
LOCUS Oryctolagus cuniculus translation initiation factor eIF2C mRNA,
DEFINITION complete cds.
ACCESSION AF005355
VERSION AF005355.1 GI:3253158
KEYWORDS
```

```
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha;
Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 3599)
AUTHORS Zou, C., Zhang, Z., Wu, S. and Osterman, J.C.
TITLE Molecular cloning and characterization of a rabbit eIF2C protein
JOURNAL Gene 211 (2), 187-194 (1998)
PUBMED 9602122
REFERENCE 2 (bases 1 to 3599)
AUTHORS Zou, C., Zhang, Z., Wu, S. and Osterman, J.C.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-1997) Chemistry, University of Nebraska, 760
Hamilton Hall, Lincoln, NE 68588, USA
FEATURES
source
1..3599
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/db_xref="taxon:9986"
/tissue_type="liver"
82..2523
/codon_start=1
/product="translation initiation factor eIF2C"
/protein_id="AAC24323.1"
/db_xref="GI:3253159"
/translation="MDIPKIDIIHYELDIKEKPPRVNREIVEHVOHFKAQIIGDR
KPEVDGRKNTYAMPPIPIGRIBXLEVLTPGEGKDIRIFKVISIKWVSCVSLQALHDALS
GRLESPVPETIQALDVMNRHLPFSMRYTPVGRSFFTAEECCNSPLGGREHVFQFHSV
RPSWLMMLNIDVSATPAQVIEFCEVLDFKSIIEQQKPLTDSORVKFTKIKG
LKVETHKQSKRYVCNVRPASHOTFPFLESQGTVECTVAQYKPKRHLVRLY
PHELCLOVQSQKTYLPLEVCNIVAGORCIKKLTNDOTSTMIRATASAPDQBEIS
KLMSASNTDPYREFGIMVKDEMTDVTGRVLPSPILYGRNKAIATPVQGVNDR
NKQHTGLEIKHAIACAPAPQCCTEVEHLKSFTEGRKISRDAQMPIQGQPCFKYAQ
GADSVGMRHNLNTAYAGLVVILPGKTPIYAKRVGDTVLGMAQCVQMKNVQR
TTPQTLNCLKINLVGNVNIILPQGRPVFPVQOPVPLGADVTHPAGDGKPSIA
AVVGSMDAHPNRYCATVRVQOHROIQDLAAMVRELLIQFYKTRFKPTRIFVRDG
VSEGOFOQLVHLLAIREACTKLEKDYQPGITFIVOKRHHTRFLCTDKNERVQKSG
NIPAGTVDTKLTHPTPEDFYLCSHNGIQGTSRPSHVLWMDNDFSSDELQILTYQL
CHTVRCRVSISIPATPAYIHLVAFRARIHLVDKEDHSAEGSHTSQSGNDRHQALAK
AVQHQDTLRTMYFA"
ORIGIN
Alignment Scores:
Pred. No.: 6,958-55 Length: 3599
Score: 565.00 Matches: 235
Percent Similarity: 40.2% Conservative: 158
Best Local Similarity: 24.1% Mismatches: 371
Query Match: 10.6% Indels: 214
DB: 14 Gaps: 38
US-10-645-746-3 (1-1020) x AF005355 (1-3599)
QY 38 LysLysValLeuLeuValAsnTrpPheLysPheSerSerLysLysIleTyrAspArgGlu 57
Db 49 AGAACATCAAGTTCAGCGCCCAATTTCTTCGAGATGGACATCCCAAAATTCACATCTAT 108
QY 58 TyrTyrGluTyrGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro 77
Db 109 CACTATGAATTCGATATCAAGCCCGAGAAA----- 138
QY 78 PheProLysLysThrGluIleProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
Db 139 TGCCCCAGAGAGTGAACAGGAGATAGTAGACACATG-----GTG 180
QY 98 ArgHisGluLysLysGlnThrAspPheIleLeuGluAspTyr-----ValPheAspGlu 115
Db 181 CAGCACTTTAAAGCCAG-----ATCTTCGCGACCGGAGCGGTGTTCGACGGG 231
QY 116 LysAspThrValTyrSerValCysArgLeuAsnThrValThrSerLysMetLeuValSer 135
Db 232 AGGAAGAACCTGTATACACGGCGATCGCTGCCCCATCGGAGGAGGAAGGTG----- 282
```


QY	136	GluLysValValLysLysAspSerGluLysLysAspGluLysAspLeuGluLysLysIle	151
DB	283	 GAGCTGGAGGTCA CGTGC GCGGAGAAGCAAGGACCGC-----ATCTTCAGGTGTCC	336
QY	156	LeuTyThrMetIleLeuThrTyArgLysLysPheHisLeuAsnPheSerArgGluAsn	175
DB	337	 ATCAAAGTG GTGTGCGTGAGCCTGCGAGCGCGCTGCACGACGCACCTTTCCGGGCGGTG	396
QY	176	ProGluLysAspGluAlaAenArgSerTyLysPheLeuLysAsnValMetThrGln	195
DB	397	 CCCAGCGTCCC GTTCCGAG-----ACGATCCAGGCCCTGGAGCTGGTCATGAGGCAT	447
QY	196	-----LysValArgTyThrAlaProPheValAsnGluGluLysValGlnPheAlaLys	213
DB	448	 TTGCCATCATGAGTACACGCTGTG-----	474
QY	214	AsnPheValTyAspAsnAsnSerIleLeuArgValProGluSerPheHisnProAsn	233
DB	475	 -----GGCGCTCTCTTCTTCACC GCGTCCGAGGCGTGTCTCCAACCT---	516
QY	234	ArgPheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyrlle	253
DB	517	 -----CTTCGGGGCGCGGAAGTGTGTCTCGCTTC CACGAC	555
QY	254	GlyIleLysGluLeuPheAspGlyGluProValLeuAsnPheAlaIleValAspLysLeu	273
DB	556	 TCTGTCCGACCTTCTCTT---TGGAAAATGATGTGAATATTGACGTGTCACGACGCGCA	612
QY	274	PheTyraenAlaProLysMetSerLeuLeuAspTyThrLeuLeuLysValAspProGln	293
DB	613	 TTCTACAAGGCG-----CAGCCAGTAATCGAGTTTGTGTGAAGTTCTGGATTTAAA	666
QY	294	SerCyAsnAspValArgLysAspLeuLysThrLysLeuMetAlaGlyLysMetThr	313
DB	667	 AGCATTGAAGAACAA-----CAGAAACCT	690
QY	314	IleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsnLeuLys---	332
DB	691	 CTGACAGATTCCCAAAGGTC AAGTTTACC AAAGAAATCAA AAGGTCTCAAGGTGGAGATA	750
QY	333	-----CysAlaGluValTrpAspAsn-----GluMetSerArgLeuThrGlu	346
DB	751	 ACGCAC TGTGGCGAGATGAAGAGGAAGTACCGCTCTGCAACGTGACCCGCGCGCCGCC	810
QY	347	ArgHisLeuThrPheLeuAspLeuCysGluGluGlnAsnSerLeuValTyThrLysValThrGly	366
DB	811	 AGCCACCGACGTTCTCCTCTGCACGACAGAGCGGGGAGACGCGTGGAGTGCAGC---	864
QY	367	LysSerAspArgGlyArgAsnAlaLysTyThrAspThrThrLeuPheLysIleTyxGlu	386
DB	865	 -----GTGGCCCCAGTAC-----TTCAAG-----GAC	885
QY	387	GluAsnLysLysPheIleGluPheProHisLeuProLeuValLysValLysSerGlyAla	406
DB	886	 AGGCACAAGCTGTTCTGCGCTACCC CACCTCCCGCTCTGCAAGTGGGACAGAGCAG	945
QY	407	LysGluTyThrAlaValProMetGluHisLeuGluValHisGluLysProGlnArgTyLys	426
DB	946	 AAGCACACTTACCTTCCCTCGAGGTCTGCAACATATGTCGCGCGGACAGAGATGATATAA	1005
QY	427	AsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThr---	445
DB	1006	 AAATTAACGGAC-----AATCAAACCTCAACTATGATCATCAGCGACCGCAGTCTGCG	1059
QY	446	HisAspTyLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGluGlu	465
DB	1060	 CCCGATCGCAAGAAGAGATCAGCAAGTTGATGTGGAACGCGCGCTTCAACACAGATCCG	1119
QY	466	LeuAsnPheValGluArgPheGlyLeuCySerLysLeuGlnMetIleGluCyProGly	485
DB	1120	 -----TACGTCCGAGTTCGGAATCAATGTGTGAAGACGAGATGACGCGTGCACGGG	1173
QY	486	LysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThrPro	505

1174	QY	CGGGTCCTGCAGCGCCCTCCATCTCTACGGGGCGCGAATGAAGCGATCGCCACGCC	1233
506	Val	-----IleArgGlyPheGlnGlnLysGlnLeuAsnVal---ValProGluLysGlu	522
1234	DB	GTGCAGGGCGTGTGGGACATGAGGACACAGAGTTCCACAGGCGATCGAGATCAGGTG	1293
523	QY	LeuCysCysAlaValPheValValAsnGlnThrAlaGlyAsnProCysLeuGluGluAsn	542
1294	DB	TGGGCCCATCGCTCTTCGGCGCCCGCCAG-----CGCCAGTGCACGGAAGTCCAC	1341
543	QY	AspValValLysPheTyrThrGluLeuLeuLeuGly-----GlyCysLys	556
1342	DB	-----CTTAAGTCTTTACAGACAGCTCAGGAAGATCTCGAGAGACGCGGCGATCGCG	1395
556	QY	sPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyValaGlnSerIle-----Me	574
1396	DB	ATCCAGGGCCAGCGCTCTTCGCAAGTACGCG-CAGGGGGCCGACAGCGTGGGGCCCAT	1454
574	QY	tTyrAspAlaThrLysAsnGlnTyrAlaPheTyrLysAsnCysThrLeuAsnThrGlyI	594
1455	DB	GTTCACAGACCTCGAAGAACACGTCAGCGC-----1482	
594	QY	eGlyArgPheGlnIleAlaThrGluAlaLysAsnMetPheGluArgLeuProAspLys	614
1482	DB	-----1482	
614	QY	sGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTyrGlyPhe	634
1483	DB	-----GGCGTCAGCTGTGTGTGATCTCTCGCGGAGACGCGCTACGCTGA	1535
634	QY	evalLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSerGluTh	654
1536	DB	AGTGAACGCGCTCGGGACACGCTGTGGGCATGGCCACGCACTGCGTCGAGATGAAGA	1595
654	QY	rValThrLysAlaLeuAsnSerLeuArgHisGlnLysGlySerLysArgIlePheTyrGl	674
1596	DB	CGTG-----CAGAGGACCAACGCGCAGACGCTGTCCAA	1628
674	QY	nIleAlaLeuLysIleAsnAlaLysLeuGlyIleAsnGlnGluLeuAspTrpSerGl	694
1629	DB	CCTCTGCTGAAGATCAACGCTGAAGCTGGCGCGGTGAACAACATCTCTG-----	1677
694	QY	uIleAlaGluIleSerProGluGluLysGluArgArgLysThrMetProLeuThrMetTy	714
1678	DB	-----CTCCGCGAGGCGAGGCCCGCGGTTCACAGACCGCGTCT---ATATT	1721
714	QY	rValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIleAlaAl	734
1722	DB	TCTGGGTGCAGATGTCACTCACCCCGCTGGGATGGCAGAACCCCTCGATCGCTGC	1781
734	QY	aValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValThrGlnGl	754
1782	DB	CGTCGTGGGAGATGGAGCGCCACCCACCGCTACTCGCCACCGCTGGGTGCAGCA	1841
754	QY	uGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIleLeuGl	774
1842	DB	G-----CACCAGGAGATCATTCAGACCTGGC	1871
774	QY	uAlaLysPheValLysLeuLeuArgGluPheAlaGluAsnAsnAspAsnArgAlaProAl	794
1872	DB	CGCCATGGTGGCGAGCTGCTCATCCAGTTCTACAAGTCCACGCGCTTCAAG---CCCAC	1928
794	QY	aHisIleValValTyrArgAspGlyValSerAspSerGluMetLeuArgValSerHisAs	814
1929	DB	GGCATCATCTTCTACCGCGAGCGCGTCTCCAGGGGCCAGTTCCACAGAGTTCTCCACCA	1988
814	QY	pGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGluArgAspGlyGluAs	834
1989	DB	CGAGCTGCTCGCCATCCGGAGGCGCTGCATCAAGCTG-----GAGNAGGA	2033
834	QY	pProGluProLysTyrThrPheIleValIleGlnLysArgHisLeuAsnThrArgLeuAr	854

```

Db      2034 CTACACCGCGGGATCACCTTCATCGTGTGCAGAGCGGCACACACCGAGCTCTTC--- 2091
Qy      854 gArgMetGluLysAsePlyeProValValaAenLysAsePLeuThrProAlaGluThrAsePva 874
Db      2092 -TGCACGGACAGACAGCGGGTCCGCAAGAGCGCGGAACATTCACAGC----- 2139
Qy      874 lalaValAlaAlaValLysGlnTrpGluGluAsePMetLysGluSerLysGluThrGlyI 894
Db      2139 ----- 2139
Qy      894 eValaAenProSerSerGlyThrThrValaAsePlyeLeuIleValSerLysThyLysPheAs 914
Db      2140 -----GSCACGACCGTGCACACGAGATCACACACCGCAGGTTTGA 2183
Qy      914 pPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgProGlyHisThyThrVa 934
Db      2184 CTTTACTACCTGTGCAGCCAGCGCGCATCCAGGGGACACAGCGCCCTCCCACTTACCACGT 2243
Qy      934 lMetTyAsePlyeGlyMetSerGlnAsePLeuValTyLysMetThrTyLysGluLeuAl 954
Db      2244 CTTGTGGATGACATCGCTCTCTCCGATGAGCTACAGATCCTGACCTACAGCTGTG 2303
Qy      954 aPheLeuSerAlaAryCArgLysProIleSerLeuProValProValHisThyAlaHi 974
Db      2304 TCACACCTACGTGTGCGGCTCCGCTCTCTATCCCGCGCGGCTTACTAGCTCA 2363
Qy      974 sLeuSerCysGluLysAlaLysGluLeuTyArgThrTyLysGluHis 990
Db      2364 CTTGTGTGCGCTTCCGCGCCAGG---TACCACCTGTGTGGAGGAACAC 2409

RESULT 14
LOCUS      BD156502              3996 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD156502
VERSION        BD156502.1 GI:27862260
KEYWORDS       JP 2002191363-A/11345.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.
REFERENCE      1 (bases 1 to 3996)
AUTHORS       Ota T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
               Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE         Primer for synthesizing full-length cDNA and use thereof
JOURNAL       Patent: JP 2002191363-A 11345 09-JUL-2002;
               HELIX RESEARCH INSTITUTE
COMMENT       OS Homo sapiens (human)
               PN JP 2002191363-A/11345
               PD 09-JUL-2002
               PF 28-JUL-2000 JP 2000280990
               PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
               PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
               PI KEIICHI NAGAI, TETSUJI OTSUKI
               PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
               10,
               PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
               Primer for synthesizing full-length cDNA and use thereof FH Key
               Location/Qualifiers
               (297)..(2642).
FEATURES             source
             1..3996
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
ORIGIN
Alignment Scores:      8.11e-55      Length:      3996
Pred. No.:             565.00      Matches:      230
Score:

```

```

Percent Similarity:      40.6%      Conservative:      170
Beat Local Similarity:   23.3%      Mismatches:       363
Query Match:            10.6%      Indels:          224
DB:                      2         Gaps:            37

US-10-645-746-3 (1-1020) x BD156502 (1-3996)

Qy      38 LysLysValLeuLeuLeuValAsnTrpPheLysPheSerSerLysIleTyRAspArgGlu 57
Db      171 AAACCAATCAAGCTCTCGCCAAATTACTTTGAGGTGGACATCCCTTAAGATCGACGTGTAC 230
Qy      58 TyrTyRGluTyRGluValLysMetThrLysGluValLeuAsnArgLysProGlyLysPro 77
Db      231 CACTACAGAGGTGACATCAAGCCGGATAAG----- 260
Qy      78 PheProLysLysThrGluLeuProIleProAspArgAlaLysLeuPheTrpGlnHisLeu 97
Db      261 TGTCCCGGTAGAGTCAACCGGGAAGTGGTGGAA-----TACATGTGTC 302
Qy      98 ArgHisGluLysGlnThrAspPheIleLeuGluAsePtyr-----ValPheAsePglu 115
Db      303 CAGCATTTCAAGCCCTCAG-----ATCTTTGGTGATCGCAAGCCTGTGTATATGA 353
Qy      116 LysAsePThrValTyRSerValCysArgLeuAsnThrValThrSerLysMetLeuValSer 135
Db      354 AAGAAGAACAATTAC-----ACTGTCACAGCACTGCCCATTTGGCAAC 395
Qy      136 GluLysVal-----ValLysLysAsePserGluLysLysAsePgluLysAsePglu 152
Db      396 GAACGGGTGCGACTTTGAGGTGACAATCCCTGGGGAAGGAAGGAT---CGAATCTTTAAG 452
Qy      153 LysLysIleLeuTyRThrMetIleLeuThrTyRArgLysLysPheHisLeuAsnPheSer 172
Db      453 GTCTCCATCAAGTGTGCTAGCCATTGTGAGCTGGCGAATGCTGCATGAGGCCCTGTGTGTCAGC 512
Qy      173 ArgGluAsePProGluLysAsePgluGluAlaAseArgSerTyLysPheLysLysAseVal 192
Db      513 GGCCAGATCCCTGTTCCCTTGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGCCACCTG 572
Qy      193 MetThrGlnLysValArgTyRAlaProPheValAseGlnGluLysValGlnPheAla 212
Db      573 GCATCC-----ATGAGGTACACCCCTGTG----- 596
Qy      213 LysAsePheValTyRAspAseAsnSerIleLeuArgValProGluSerPheHisAsePPro 232
Db      597 -----GGCCGCTCTCTTCTCTCACCCCTGAGGGCTACTACCACCG 638
Qy      233 AsnArgPheGluGlnSerLeuGluValAlaProArgIleGluAlaTrpPheGlyIleTyR 252
Db      639 -----CTGGGGGTGGCGCGAGGTCTGTGTCGCTTTCAC 674
Qy      253 IleGlyIleLysGluLeuPheAsePgluProValLeuAsePheAlaIleValAsePlye 272
Db      675 CAGTCTGTGGCCCTGCCCATG---TGGAGATGATGCTCAACATGATGTCAGCCACT 731
Qy      273 LeuPheTyRAsnAlaProLysMetSerLeuLeuAsePtyRLeuLeuIleValAsePPro 292
Db      732 GCCTTTTATAAGSCA-----CAGCCAGTGTAGTTCATGTGTGAGGTGTGGACATC 785
Qy      293 GlnSerCysAseAspAsePValArgLysAsePLeuLysThrLysLeuMetAlaGlyLysMet 312
Db      786 AGGAACATAGATGACAGCCCAAG----- 809
Qy      313 ThrIleArgGlnAlaAlaArgProArgIleArgGlnLeuLeuGluAsePLeuLys 332
Db      810 CCCCTCAGGACTCTCAGCGCGTTCGCTTACCAGGAGATCAAGGCCCTCAAGGTGGA 869
Qy      333 -----CysAlaGluValTrpAsePhe-----GluMetSerArgLeuThr 345
Db      870 GTCACCCACTGTGGACAGATGAAGAGAGTACCGCGTGTGTAAATTACCGCTCGCCCT 929
Qy      346 GluArgHisLeuThrPheLeuAsePLeuCysGluGluAsePLeuValTyRValThr 365

```

```
Db 930 GCTAGCCATCAGACATTCCCTTACAGCTGGAGAGTGACAGACTGTG----- 977
Qy 366 GlyLysSerAspArgGlyArgAsnAlaLysLysTyrAspThrLeuPheLysLeuTyr 385
Db 978 -----GAGTGCACAGTGGCAGCATTTTC 1001
Qy 386 GluGluAsnLysLysPhe-----IleGluPheProHisLeuProLeuValLysValLysSer 404
Db 1002 AAGCAGAAATATAACCTTCAGCTCAAGTATCCCATCTGCCCTCCCTACAAAGTTGGCCAG 1061
Qy 405 GlyAlaLysGluTyrAlaValProMetGluHisLeuGluValHisGluLysProGluArg 424
Db 1062 GAACAAAGCATACCTACCTTCCTCCTAGAGCTGTGAACATT---GTGGCTGGCGCAGGC 1118
Qy 425 TyrLysAsnArgIleAspLeuValMetGlnAspLysPheLeuLysArgAlaThrArgLys 444
Db 1119 TGTATTAAAGCTGACGCGCAGCAACACAGACCTCGACCATGATGAAGGCCACAGCTAGATCC 1178
Qy 445 ProHisAspTyrLysGluAsnThrLeuLysMetLeuLysGluLeuAspPheSerSerGlu 464
Db 1179 GCTCCAGACAGACAGGAGGAGATCAGTCCCTGATGAAGAAATGCCAGCTCAACCTTAGAT 1238
Qy 465 GluLeuAsnPheValGluArgPheGlyLeuCysSerLysLeuGlnMetIleGluCysPro 484
Db 1239 CCC-----TACATCCAGGAATTGGGATCAAGTGAAGTGCATGACGAGGTGACA 1292
Qy 485 GlyLysValLeuLysGluProMetLeuValAsnSerValAsnGluGlnIleLysMetThr 504
Db 1293 GGGCGAGTGTCTCGCGCGCCCATCTTGCAGTACGCGCGCGCAACCGGCCATTGCCACA 1352
Qy 505 Pro-----ValIleArgGlyPheGlnGluLysGlnLeu---AsnValVal 518
Db 1353 CCCAATCAGGGTGTCTGGACATCGCGGGG-----AAACAGTTCTCAATGGGATT 1403
Qy 519 ProGluLysGluLeuCysCysAlaValPheValValAsnGluThrAlaGlyAsnProCys 538
Db 1404 GAGATCAAGTCTGGCCCATCGCTGCTTCGCCACCCCAAAA-----CAGTGT 1451
Qy 539 LeuGluGluAsnAspValValLysPheTyrThrGluLeuIle-Gly----- 553
Db 1452 CGAAGAG-----GTGCTCAAGAACTTCACAGACCAGCTGCGGAAGATTTCCAGGAT 1505
Qy 554 ----GlyCysLysPheArgGlyIleArgIleGlyAlaAsnGluAsnArgGlyAlaGlnSe 572
Db 1506 GCGGGATGCTATCCAGGGTCAACCTTGTTCGCAAAATATGCACAGGGGGCA-GACAG 1564
Qy 572 rIleMetTyrAspAlaThrLysAsnGluTyrAlaPheTyrLysAsnCysThrLeuAsnTh 592
Db 1565 CGTG----- 1568
Qy 592 rGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAsnMetPheGluArgLeuPr 612
Db 1569 -----GAGCCTATGTTCCGGCATCTCAA 1591
Qy 612 oAspLysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAsnAlaTy 632
Db 1592 GAACACCTACTAGGGCTGCGAGTCACTATTGTCATCTCCAGGAAGACGCGGTGTA 1651
Qy 632 rGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAsnGlnHisIleThrSe 652
Db 1652 TGCTGAGGTGAACGTGCGGATACACTCTTGGGAAATGGCTACCGAGTGTGTGAGGT 1711
Qy 652 rGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePh 672
Db 1712 GAAGAACGTGGTCAAGACC-----TCACCTCAGACTCT 1744
Qy 672 eTyrGlnIleAlaLeuLysIleAsnAlaLysLeuGlyGlyIleAsnGlnLeuLeuAspTr 692
Db 1745 GTCCAACTCTGCTCAAGATCAATGTCAAACTTGGTGGCAATTAACAGCATCTCA----- 1799
Qy 692 pSerGluIleAlaGluIleSerProGluLysGluArgArgLysThrMetProLeuTh 712
Db 1800 -----GTCCCAACACGAGCTCTGCGGTTTTTCAACAGCAGCATG--- 1838
```

```
Qy 712 rMetTyrValGlyIleAspValThrHisProThrSerTyrSerSerGlyIleAspTyrSerIle 732
Db 1839 -ATATTCTGGAGCAGATGTTACACACCCCGCAGAGGGATGGGAAAAAACCTTCTAT 1897
Qy 732 eAlaAlaValValAlaSerIleAsnProGlyGlyThrIleTyrArgAsnMetIleValTh 752
Db 1898 CACAGCAGTGGTAGGAGTAGTATGGATGCCACCCCGAGCGCATCTGTGTACTGTGCGGT 1957
Qy 752 rGlnGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIle 772
Db 1958 ACAG-----COACACCGCGCAAGAT 1978
Qy 772 eLeuGlu-----AlaLysPheValLys---LeuLeuArgGluPheAlaGluAsnAsnAs 789
Db 1979 CATTGAAGACTTGTCTTACATGCTGCTGAGCTCTCATCAATTCATCAAGTCCACCG 2038
Qy 789 pAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLe 809
Db 2039 TTTCAAG---CCTACCCGCAATCATCTTCTCCGAGATGGGGTGCCTGAAGCCAGCTACC 2095
Qy 809 uArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGl 829
Db 2096 CCAGATACTCCATATGAGCTACTGGCCATTGCTGATGCTGCATCAAACTG----- 2147
Qy 829 uArgAspGlyGluAspProGluProLysTyrThrPheIleValIleGlnLysArgHisAs 849
Db 2148 -----GAAAGAGACTTACCAGCTGGGATCACTTATATTGTGTGTCAGAAACCCATCA 2200
Qy 849 nThrArgLeuLeuArgArgMetGluLysAspLysProValValAsnLysAspLeuThrPr 869
Db 2201 CACCCGCTTTTC---TGCTGCACAGAAATGGCGGAATGGGAAGAGTGGTAACATCCC 2257
Qy 869 oAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLysGluSe 889
Db 2258 AGCT----- 2261
Qy 889 rLysGluThrGlyIleValAsnProSerSerGlyThrThrValAspLysLeuIleValSe 909
Db 2262 -----GGACCAACAGTGGACCAACATCACCCA 2290
Qy 909 rLysTyrLysPheAspPheLeuAlaSerHisHisGlyValLeuGlyThrSerArgPr 929
Db 2291 CCCATTGATTTGACTTCTATCTGTGCACCCAGCAGGATCCAGGCGCACCGAC 2350
Qy 929 oGlyHisTyrThrValMetTyrAspLysGlyMetSerGluAspGluValTyrLysMe 949
Db 2351 ATCCCATTTACTATTGTTTGGATGACAAACCGTTTCACAGCAGATGAGCTCCAGATCCT 2410
Qy 949 tThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysPheProIleSerLeuProValPr 969
Db 2411 GAGCTACCACTGTGCCACACTTACGTACGATGCACACGCTCTGTCTCTATCCAGCACC 2470
Qy 969 oValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGl 989
Db 2471 TGCTTACTATGCCCGCTGTGCTGCTTCCGGGCACGA---TACCACCTGTGTGACAAAGA 2527
Qy 989 uHisTyrIleGlyAsp 994
Db 2528 GCATGACAGTGGAGAG 2543
```

```
RESULT 15
AX877133
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
```

AX877133 3996 bp DNA linear PAT 17-DEC-2003
Sequence 12038 from Patent EP1074617.


```
QY 539 LeuGluGluAenAaspValValLysPheTyrThrGluLeuIle-Gly----- 553
Db 1452 CGAAGAG-----GTGCTCAAGAACTTCACAGCAGCTGCGGAAGATTTCACAGAT 1505
QY 554 ----GlyCysLysPheArgGlyIleArgIleGlyAlaAenGluAenArgGlyAlaGlnSe 572
Db 1506 GCGGGGATGCCATATCCAGGGTCAACCTTGTTCTGCAAAATATGCACAGGGGGA-GACAG 1564
QY 572 rIleMetTyrAspAlaThrLysAenGluTyrAlaPheTyrLysAenCysThrLeuAenTh 592
Db 1565 CGTG----- 1568
QY 592 rGlyIleGlyArgPheGluIleAlaAlaThrGluAlaLysAenMetPheGluArgLeuPr 612
Db 1569 -----GAGCTATGTTCGGGCATCTCAA 1591
QY 612 oAspLysGluGlnLysValLeuMetPheIleIleSerLysArgGlnLeuAenAlaTy 632
Db 1592 GNACACCTACTCAGGGCTGCAGCTCATTTATGTCTCTGCCAGGGAAGACGCCGGTGA 1651
QY 632 rGlyPheValLysHisTyrCysAspHisThrIleGlyValAlaAenGlnHisIleThrSe 652
Db 1652 TGCTGAGGTGAACGTGCGAGATACACTCTTGGGAATGGCTACGCAGTGTGTGCAGGT 1711
QY 652 rGluThrValThrLysAlaLeuAlaSerLeuArgHisGluLysGlySerLysArgIlePh 672
Db 1712 GAGAACGTGTCAAGACC-----TCACCTCAGACTCT 1744
QY 672 eTyrGlnIleAlaLeuLysIleAenAlaLysLeuGlyGlyIleAenGlnGluLeuAaspTr 692
Db 1745 GTCCAACTCTGCTCAAGATCAATGTCAAACTTGGTGGCATTTAACAGCATCTTA----- 1799
QY 692 pSerGluIleAlaGluSerProGluGluLysGluArgLysThrMetProLeuTh 712
Db 1800 -----GTCCCAACACAGCGCTCTGCGGTTTTTCAACAGCCAGTG-- 1838
QY 712 rMetTyrValGlyIleAspValThrHisProThrSerTyrSerGlyIleAspTyrSerIl 732
Db 1839 -ATATTCTGGGAGCAGATGTACACACCCCGAGCGGGATGGGAAAAAACCTTCTAT 1897
QY 732 eAlaAlaValAlaSerIleAenProGlyGlyThrIleTyrArgAenMetIleValTh 752
Db 1898 CACAGCAGTGTAGGCAGTATGATGCCACCCAGCCGATCTGTCTACTGTGCGGT 1957
QY 752 rGlnGluGluCysArgProGlyGluArgAlaValAlaHisGlyArgGluArgThrAspIl 772
Db 1958 ACAG-----CGACCACGCGCAAGAGAT 1978
QY 772 eLeuGlu-----AlaLysPheValLys---LeuLeuArgGluPheAlaGluAenAenAs 789
Db 1979 CATTGAAGACTTGTCTACATGGTGGCGTGAGCTCTCTCATCAATTCTACAAGTCCACCCG 2038
QY 789 pAsnArgAlaProAlaHisIleValValTyrArgAspGlyValSerAspSerGluMetLe 809
Db 2039 TTTCAG--CCTACCGCATCATCTTCTCGCAGATGGGTGCTTGAAGGCCAGCTACC 2095
QY 809 uArgValSerHisAspGluLeuArgSerLeuLysSerGluValLysGlnPheMetSerGl 829
Db 2096 CCAGATACTCCACTATGAGCTACTGGCCATTCTGTGATCGCTGCATCAAACTG----- 2147
QY 829 uArgAspGlyGluAspProGluProlLysTyrThrPheIleValIleGlnLysArgHisAs 849
Db 2148 -----GAAAAGACTTACCGAGCTGGGATCATCTTATTTGTGTGCGAAGAACGCCATCA 2200
QY 849 nThrArgLeuLeuArgArgMetGluLysAspLysProValValAenLysAspLeuThrPr 869
Db 2201 CACCGGCTTTTC---TGTCTGACAGAATGGCGAATTTGGGAAGAGTGTAAACATCCC 2257
QY 869 oAlaGluThrAspValAlaValAlaValLysGlnTrpGluGluAspMetLysGluSe 889
Db 2258 AGCT----- 2261
QY 889 rLysGluThrGlyIleValAenProSerSerGlyThrThrValAspLysLeuIleValSe 909
```

```
Db 2262 -----GGGACCACAGTGGACACCAACATCACCCA 2290
QY 909 rLysTyrLysPheAspPhePheLeuAlaSerHisGlyValLeuGlyThrSerArgPr 929
Db 2291 CCCATTTGAGTTTGACTTCTATCTGTCGCCACCGCAGGCATCCAGGSCACAGCCGACC 2350
QY 929 oGlyHisTyrThrValMetTyrAspAspLysGlyMetSerGlnAaspGluValTyrLysMe 949
Db 2351 ATCCCATTTACTATCTTTTGGGATGACAAACCGTTTCACAGCAGATGAGCTCCAGATCCT 2410
QY 949 tThrTyrGlyLeuAlaPheLeuSerAlaArgCysArgLysProIleSerLeuProValPr 969
Db 2411 GAGGTACACAGCTGTCACACTTACGTACGATGCACACGCTCTCTCTATCCAGCAC 2470
QY 969 oValHisTyrAlaHisLeuSerCysGluLysAlaLysGluLeuTyrArgThrTyrLysGl 989
Db 2471 TGCTTACTATGCCCGCTTGGTGGCTTTCCGGGCACGA---TACCACCTGGTGGACAGGA 2527
QY 989 uHisTyrIleGlyAasp 994
Db 2528 GCATGACAGTGGAGAG 2543
```

Search completed: July 6, 2006, 02:08:28
Job time : 12644 secs

THIS PAGE BLANK (except)